This dissertation has been
microfilmed exactly as received 66-4138

BEAUCHAMP, James Warren, 1937-
ELECTRONIC INSTRUMENTATION FOR THE
SYNTHESIS, CONTROL, AND ANALYSIS OF
HARMONIC MUSICAL TONES.

University of Illinois, Ph.D., 1965
Engineering, electrical

University Microfilms, Inc., Ann Arbor, Michigan

ELECTRONIC INSTRUMENTATION FOR THE SYNTHESIS, CONTROL, AND ANALYSIS OF HARMONIC MUSICAL TONES

BY

JAMES WARREN BEAUCHAMP
B.S., University of Michigan, 1960
M.S., University of Michigan, 1961

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois, 1965

Urbana, Illinois

UNIVERSITY OF ILLINOIS

THE GRADUATE COLLEGE

September, 1965

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY

SUPERVISION BY James Warren Beauchamp

ENTITLED Electronic Instrumentation for the Synthesis, Control, and Analysis of Harmonic Musical Tones

BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF Doctor of Philosophy in Electrical Engineering

Lejaren A. Hiller Jr.
In Charge of Thesis

E.C. Jordan
Head of Department

Recommendation concurred in†

Lejaren A. Hiller Jr.
C. W. Mendel
M E Van Valkenburg
E.C. Jordan
M L Babcock

Committee
on
Final Examination†

† Required for doctor's degree but not for master's

D517

ACKNOWLEDGEMENT

I am deeply grateful to Professor Lejaren A. Hiller, Jr., my thesis advisor, for his assistance, advice, and encouragement throughout the course of this research. To the Magnavox Company I wish to express gratitude for creating the Magnavox Fellowship in Electrical Engineering and Music, which has made my work on this project possible during the past three years. I am very thankful to both Professor Hiller and Professor E. C. Jordan, Head of the Department of Electrical Engineering, for their support of my work and for their time spent in promoting an interest between the Magnavox Company and the University of Illinois.

I am indebted to Dr. William B. Ribbens of Ann Arbor, Michigan for his interest in my work and for his suggestion of the method for separating audio harmonics by means of ultrasonic frequencies. Also, I would like to thank Professors Salvatore Martirano, Kenneth Gaburo, and Herbert Brün of the School of Music, for many inspiring conversations and useful suggestions. They have created magnetic tape compositions which demonstrate the sound generating capabilities of the equipment described herein.

Special thanks are due Mr. Bela Sandorfi and Mrs. Susan Torrico for preparing the illustrations.

Finally, I want to thank my wife, Elizabeth, who has been a source of unceasing confidence and has assisted in the preparation of this thesis.

TABLE OF CONTENTS

CHAPTER 1 THE IMPORTANCES OF HARMONIC TONES AND METHODS FOR THEIR SYNTHESIS AND CONTROL 1

1.1 Introduction 1
1.2 Single Sounds and Harmonic Tones in Music 5
1.3 The Complexity of Musically "Rich" Harmonic Tones 7
1.4 Synthesis of Harmonic Tones 11
1.5 Voltage Control of Determinative Parameters 12
1.6 Some Applications of Harmonic Tone Synthesis 14

CHAPTER 2 DESIGN OF A HARMONIC TONE GENERATOR 17

2.1 Introduction 17
2.2 Operational Form 17
2.3 Three Possible Methods for Additive Generation of Harmonic Tones 18
  a. Hetrodyne Method 19
  b. Multiplier-Subtracter Method 24
  c. Voltage-Controlled Oscillator Method 25
2.4 Some Design Considerations 27
  a. Number of Harmonics 27
  b. Control of the Fundamental Frequency 29
  c. Harmonic Amplitude Control 32
  d. Harmonic Relative Phase Control 34
  e. Purity of the Output Signal 35
2 5 Description of Circuitry for a Practical Additive Synthesis Device, the Harmonic Tone Generator 36

CHAPTER 3 CIRCUITRY FOR THE GENERATION AND CONTROL OF PARAMETER VOLTAGES 51

3.1 Introduction 51
3.2 Time-Envelope and Gate Control 54
3.3 Keyboard Pitch and Rhythm Control 60
3.4 Rhythm Control by Tape and AC-DC Converter 63
3.5 Analog Multiplier 65
  a. A Practical Phase Shift Multiplier 66
  b. Parameter Control by Multiplication 73
  c. Other Uses for Multipliers in Electronic Music 74
  d. Some Other Possible Designs for an Electronic Multiplier 75

CHAPTER 4 VARIABLE FORMANT FILTER (PROGRAM EQUALIZER), THEORY AND DESIGN 79

4.1 Introduction 79
4 2 A "Best" Criterion for Formant Curve Simulation 82

4.3 Computation of Ripple and Notch Characteristics of an Infinite Set of Bandpass Filters 83
4.4 Calculations for Finite Sets of Bandpass Filters 95
4.5 A Practical Variable Formant Filter 102
4.6 Computer Simulation of Variable Formant Filters 105

CHAPTER 5 FOURIER SERIES METHODS FOR ANALYSIS OF A TRANSIENT HARMONIC TONE AND A PROJECTED ANALOG SYSTEM FOR ANALYSIS AND RESYNTHESIS OF MUSICAL TONES 107

5.1 Introduction 107
5.2 Fourier Series Representation of a Quasi-Periodic Signal 109
5.3 Frequency Domain Equivalent of Fourier Series Coefficient Calculation 112
5.4 Rectangular Filter Approach 115
5.5 Comparison of Three Fourier Series Representations of a Simple Test Signal 118
5.6 A Projected Analog System for Analysis and Resynthesis of Musical Tones 127

CHAPTER 6 A PERCEPTUAL CHARACTERISTIC OF HARMONIC TONES: THRESHOLD FOR "HEARING OUT" THE UPPERMOST PARTIAL 134

6.1 Introduction 134
6.2 Experimental Method 135
6.3 Calculations 138
6.4 Conclusions 138
6.5 The Listeners 140

BIBLIOGRAPHY 142

VITA 144

Chapter 1

THE IMPORTANCE OF HARMONIC TONES AND METHODS FOR THEIR SYNTHESIS AND CONTROL

1.1 Introduction

During the past three years, I have investigated the use of certain new and flexible forms of analog circuitry for the generation of sound materials to be used in electronic music, for electronic musical instruments, and for studies in musical sound perception. In this investigation I have stressed the synthesis and control of the timbre of harmonic tones (tones which consist of harmonic partials). Several special devices which were constructed for this investigation are the following:

(1) The Harmonic Tone Generator. This electronic device employs additive synthesis to produce at its output a harmonic tone with its fundamental frequency and harmonic amplitudes independently variable according to the values of several input parameter control voltages. The design of this device is discussed in Chapter 2.

(2) Six Time-Envelope Networks. A voltage step at the input of one of these networks produces a time-varying DC voltage output which is used primarily for building up attack and decay envelopes for musical sounds.

(3) Six Time-Delay Gate Networks. A voltage step at the input of one of these networks produces another voltage step at the output which is time-delayed with respect to the input by a pre-set amount (variable in three ranges from 0 to 5 seconds). However, the output returns to zero when the input does. Also, by

means of a separate part of the circuit an external signal can be turned on and off in synchronism with the output.

(4) Keyboard. A keyboard was wired with a set of variable resistors in conjunction with a reference voltage which produces a pitch-control voltage corresponding to the key depressed.

(5) AC-DC Converter. A high frequency tone burst (5KC or greater) at the input of this circuit produces a DC output which can be used to drive the Time-Envelope Networks or the Time-Delay Gate Networks. Hence, tone bursts recorded on magnetic tape according to a rhythmic scheme can be used to actuate these networks.

(6) Analog Multiplier. The output voltage of this circuit is proportional to the product of two input voltages which may be of either polarity. Although the multiplier was designed primarily for synthesis of special parameter control voltages, its wide frequency range (10KC) and its wide dynamic range on one input (75 db), permit it to be used as a balanced AM modulator (or "ring" modulator) or for envelope control of arbitrary signals.

Devices (2) through (6) are discussed in Chapter 3.

(7) Variable Formant Filter. The output of this filter is the sum of the attenuated outputs of twelve band pass filters with band-centers logarithmically spaced between 189 cps and 2400 cps. The transfer response of this filter is capable of assuming a large number of distinct shapes corresponding to the settings of twelve attenuator controls. Hence, by subtractive synthesis, a wide variety of output spectra can be developed from wide-band flat spectrum inputs such as white noise and pulse waves. Curves for the design of variable formant filters

according to a "best" criterion for up to 10 filters per octave and three basic transfer functions are derived in Chapter 4.

These devices are now installed at the University of Illinois Experimental Music Studio.[1] The overall philosophy of the design of this system of devices for generating musical sounds is similar to that applied in the design of an analog computer. First, in the case of a computer, electrical signals (to be converted into sound signals) are created in this system by means of a set of well-defined electrical generators and converters which can in turn be connected together in a very large number of possible combinations. Second, DC voltage signals which represent simple mathematical functions (a basic characteristic of analog computers) are used extensively for the specification and control of musical parameters such as amplitude, pitch, and tone color. However, in a music system such as this, the absolute phase of an oscillator output is usually indeterminate and feedback is seldom employed or required, in contrast to the ordinary use of an analog computer. In this music system, the output which is to be synthesized is specified *a priori* as an explicit function of time and of the defining parameters. On the other hand, with an analog computer the output, which is a solution to a problem, is only specified as an implicit function of these variables. Another important difference is that signal bandwidths needed for music (20,000 cps for output signals and 1,000 cps for parameter control) are generally much higher than those ordinarily required in analog computers (about 50 cps).

The use of DC control voltages permits the continuous control of important musical parameters within a single sound event. This is a very important concept which now can be exploited in electronic music; until just recently this concept has been lacking in practically all electronic tape music. However, within the past year and a half several compositions have been completed at the Experimental Music Studio which utilize the Harmonic Tone Generator and other devices in varying degrees. Among them are "Futility 1964" by Herbert Brun, "Machine Music" by Lejaren Hiller; "Underworld" by Salvatore Martirano; and "Antiphonics III", "Lemon Drops" and incidental music for "Hydrogen Jukebox" (a play) by Kenneth Gaburo. "Lemon Drops" is an electronic tape composition which was created almost entirely from sounds produced by the Harmonic Tone Generator.

---

[1] Hiller, L.A. Jr.: "An Integrated Electronic Music Console", *J. Audio Eng. Soc.*, *13(2)*, 142-150, 1965.

In Chapter 5, I indicate how the electronic method of additive harmonic synthesis, which utilizes a carrier frequency technique, can be extended for the analysis of a sound in terms of the parameters required to define it as a harmonic tone. Three mathematically simple versions of harmonic analysis by time-variant Fourier series calculations are presented and compared. These results are applicable to the electronic analysis system, since this system is an analog of the time-variant Fourier series method.

Finally, the sound generation system, which comprises the seven devices listed above, can also be used for studies in sound perception. The results of one such study, using the Harmonic Tone Generator, are given in Chapter 6. This was an attempt to determine the intensity of a single harmonic, in the presence of the set of lower harmonics, required for the pitch of this harmonic to be perceived as distinct from the pitch of the composite harmonic tone. The results indicate that for the first six harmonics the average level required is about 20 phons (subjective intensity units) below the combined loudness of the lower harmonics, independent of fundamental frequency from 65 to 520 cps and independent of harmonic number.

1.2 Single Sounds and Harmonic Tones in Music

Music is basically an organization of sounds within a frequency domain and a time domain framework. The "message" contained in a piece of music almost always has no concrete referential meaning. Nevertheless, the composer of the music means that a sense of multifarious structure be conveyed to the listener, and that he be capable (at least subconsciously) of aurally decomposing the sound into simpler constituent parts, making a variety of associations among them, and deriving enjoyment from this process. In conventional music, these constituent parts are usually recognized as either "single sounds" or "ensemble sounds" which are composed of single sounds. At times the distinction between these two types of sounds may be quite subtle. Experience in various electronic music studios has shown that if the sound of several instruments simultaneously playing a chord is altered by removing the attack portion of the chord, the aural effect is more like that of a complex single sound. However, when all the aural clues are present, especially with respect to differences in time and space location, the musically trained person has no trouble separating all the single sounds present in a complex musical texture. In fact, with experience he can learn to construct a reasonably accurate score from repeated listenings to the music.

By elimination, we are left with the definition of a single sound as being a sound which cannot readily be decomposed by a listener into simpler parts. Sounds of this type not only include the tones of the common wind, string, and percussion musical instruments but also many sounds in nature and human speech. We are apt, due to our experience, to think of a single sound as that which emanates from a single object or spatial location. However, in the case of electronically generated sound this generalization does not hold at all since it is quite easy to reproduce the sound of a full orchestra through a single loudspeaker or, on the other hand, to split a single sound into parts and generate them from several different spatially separated loudspeakers. Characteristics which prevent the listener from aurally diagnosing the frequency spectrum of a single sound are: (1) Lack of distinction between the times of entrance of the various frequency components; (2) Small frequency separation between discrete components or no separation (continuous frequency spectrum); (3) Harmonic relationship of components.

A "tone" can be defined to be a single sound with a definite subjective pitch. This pitch generally results from the predominant concentration of spectral energy at a single position in the frequency domain or from the definiteness of a certain time period of repetition. Harmonic tones, which are composed of frequencies integrally related to some audible fundamental frequency, are of particular importance in music. Under most conditions the subjective musical pitch of a harmonic tone is with good approximation proportional to the logarithm of its fundamental frequency (whether or not the fundamental is actually present in the sound). For a given pitch the timbre is determined by the relative intensities and apparently to a much lesser degree, the phases of the constituent harmonic frequencies. The forced vibration tones of the ordinary acoustical musical instruments, executed by bowing strings and blowing into tubes, are harmonic to a considerable degree. So are the damped vibration tones of certain percussion instruments such as the piano and the guitar. For centuries, harmonic tones have been the predominant single sounds used in music of the Western world, and because of the definite scalar ordering of the pitches of these tones, and the ease of producing all audible frequencies, irrespective of timbre, it is not surprising that the theory of pitch organization has reached a high stage of development. However, since there exists no obvious ordering of timbre, and, until recently, there has been no method of systematically and independently varying the parameters which define it, theories of timbre organization are practically nonexistent. Therefore, the composer of timbre relationships has been forced to rely upon pragmatic or _a priori_ judgments, and consequently his ordering of timbres has been for the most part restricted to making "differences" among them.

1.3 The Complexity of Musically "Rich" Harmonic Tones

The question of what constitutes the timbre of a harmonic tone is even more complex than determining the effect of different relative intensities and phases of the harmonics and the fundamental frequency. There exists an overwhelming amount of evidence to testify that on the level of sophistication of sound production provided by the conventional musical instruments, the "richness" of tones depends on the variation of these parameters with respect to time. That is to say, the precise manner in which the harmonic intensities build up to their final average values at the start of a tone, the variations of these intensities during the course of the tone, and other factors such as additive noise, pitch deviation, and small amounts of frequency and amplitude modulation also play a major role in determining the timbre characteristics of a tone and give it a certain "warmth" which is lacking in less complex electronically generated tones.

Five musical terms commonly used to describe timbre are: pitch, tone quality, attack, vibrato, and tremolo. These refer technically to the fundamental frequency, the average steady-state harmonic spectrum, the start of the tone, and regularly fluctuating changes of pitch and amplitude. However, they are oversimplified descriptions of what may be very complex phenomena. For instance, the musician refers to a "soft," "dull," "brillant," or "nasal" tone quality, while being quite unaware of the type of the harmonic spectrum which produces this effect. Certainly, at a fixed pitch and to a first-order approximation tone quality must be determined by the average relative levels of the harmonics. But whether or not it is possible to perceive tone quality as a fixed parameter over a significant pitch range, and, if so, how the harmonic distribution must vary (if it must at all) with respect to pitch in order to keep tone quality constant, are questions which cannot be answered with certainty at the present time.

Nevertheless, it is a well-known fact that the tone qualities of many musical instruments depend on "formant" characteristics. These are frequency response curves usually displaying one or more definite resonances or anti-resonances. In this case, the individual levels of the harmonics change with pitch and are proportional to the corresponding values of a formant characteristic at each of the harmonic frequencies. The formant curves for many of the nonpercussive musical instrument tones have been measured in detail by Luce.[2] The tone of a bassoon is a good example of a sound which is easily recognized in terms of a simple formant characteristic, a resonance at about 500 cps. Recognition of

---

2. Luce, D.A.: "Physical Correlates of Nonpercussive Musical Instrument Tones," unpublished Doctoral Dissertation, M.I.T., 1963, pp. 225-323.

synthetic bassoon tones is enhanced if the pitch is varied, probably because the resonance region is defined more accurately by the resulting movement of the harmonics within the boundary of the formant curve. Also, many of the sustained vowel sounds of speech can be identified from the positions of two or more formant frequencies, independent of fundamental frequency.

On the other hand, a type of tone quality invariance may also exist with respect to transposition of a sound spectrum. For example, a brass tone retains a certain identity when the pitch is changed but the spectrum "shape" is held fixed. A trumpet tone, when recorded and played back at twice the original speed (a transformation invariant with respect to spectrum shape), sounds like a toy trumpet. At half speed it sounds more like a trombone and at one-quarter speed, like a tuba. High tones of the cello can in certain contexts be confused with middle register violin tones even though the cello formant characteristic is roughly a transposition of the violin characteristic by more than an octave lower. (Besides tone quality, these intra-family confusions may be due to the existence of other parameters which describe a family of musical instruments as a whole, such as peculiar amplitude and frequency changes with respect to time). The idea of transposition invariance may have some validity for the case of sustained vowel sounds. Although for the same spoken vowel the absolute values of the formant frequencies vary considerably for different people, especially when children are included in the analysis, there is reason to believe that the ratios of the first and second formant resonances of a sustained vowel are sufficiently constant to allow their identification.[3] In some sense, this part of the formant characteristic for one person may be treated as a frequency transposition of that of another person.

---

[3] Pierce, J.R. and David, E.D.: Man's World of Sound, Doubleday & Company, Garden City, N.Y., 1958, p. 98.

The transient behaviors of the attack portions of the tones of several wind and string instruments have been measured by Luce.[4] His data show that a great deal of information is required to duplicate these attack transients precisely. Nevertheless, despite the large amounts of randomness in the harmonic intensities and phases which are present in many cases, there appears to be a considerable amount of statistical order which can be used to distinguish one musical instrument from another. The attacks are certainly not describable by total envelope modulation of a basic waveform. For several instruments Luce determined that the harmonics reach full intensity at slightly different times, the time "delay" between the entrances of partials was measured to be on the order of 10 milliseconds.

LeCaine has reported attempts to simulate vibrato and tremolo effects by constant sine wave frequency and amplitude modulation and by random modulation at rates varying between 5 and 7 cps.[5] Both of these types of simulation tend to yield a monotonous sound. I would suspect that "naturalness" of vibrato can be achieved by a control of the rate and amount of modulation with respect to time as well as the control of certain statistical parameters which describe the random components. Also, the introduction of reverberation seems to "blur" monotonous repetitions and increase "naturalness".

---

4. Luce, D.A. : "Physical Correlates of Nonpercussive Musical Instrument Tones," unpublished Doctoral Dissertation, M.I.T., 1963, pp. 100-224.

5. LeCaine, H. : "Electronic Music", Proc. Inst. Rad. Eng., 44 (4), 457-478, 1956.

1.4 Synthesis of Harmonic Tones

In electronic music, two of the most useful methods for creating sounds are *additive synthesis* and *subtractive synthesis*. In additive synthesis, the total sound signal is formed by the superposition of separate constituent parts according to the proportions desired. In subtractive synthesis, a signal is generated which contains the desired frequency components, and a formant filter is used to shape the distribution of these components. More precisely, in the additive synthesis of a harmonic tone, sine waves tuned exactly in frequency ratios 1:2:3: . . . are summed together, and the desired tone quality is developed by independent adjustment of the relative amplitude of each harmonic. A more general additive harmonic tone synthesizer would have provision for changing the phase of each harmonic and the fundamental frequency independent of the other parameters. In subtractive synthesis, an easily generated waveform composed of a large number of harmonics, with amplitudes which do not diminish too rapidly as the harmonic number increases, is used as the input to a spectrum-shaping filter. In order to change the spectrum, the filter transfer response is varied. Two other methods commonly used to generate sounds, including harmonic tones, are *modulation synthesis*, whereby two simple signals are combined by amplitude or frequency modulation to form a more complex signal, and *waveform synthesis*, in which case the wave shape of the output of a generator can be varied within certain restrictions. Under the assumption that spectral distribution is the most important factor determining tone quality, these two methods do not seem to be very flexible for scientific investigations since the distributions which arise from given syntheses are not obviously related to the parameters specified.

Subtractive synthesis is the most common method of tone generation employed in electronic organs and other electronic musical instruments for several reasons

First, waveforms such as the square wave, the pulse, and the sawtooth wave are rich in harmonics and can easily be generated by electronic circuits. Second, fixed filters for shaping a number of differentiated spectrums can be economically produced. Third, compared to additive synthesis, subtractive synthesis has no audible upper limit to the number of frequencies which can be generated; therefore, complex or "nasal" tone qualities can be produced as easily as "pure" tone qualities with low order harmonic content. Fourth, formant filtering of waveforms represents a reasonable analog to the acoustic behavior of ordinary musical instruments, including the organ, which is the instrument most builders of electronic musical instruments attempt to simulate.

On the other hand, for scientific investigations additive synthesis allows much more control over the harmonic spectrum, particularly with regard to changes of this spectrum with respect to time. Also, the variety of steady state distributions which can be produced by subtractive synthesis is limited by the frequency responses of available filters, and the structure of the transient parts of the tone produced is restricted to that which results from total modulation of the steady state distribution (or filtered parts of it). These limitations do not exist for a system of additive synthesis because the harmonics are created separately, controlled separately, and then superimposed.

1.5 Voltage Control of Determinative Parameters

Only the values of the relative amplitudes and phases of the harmonics and the fundamental frequency convey any information about a harmonic tone. In music these values must be continually changing if there is to be any information flow with respect to time. Since control by manual operation always involves restrictions on speed and accuracy, a more fruitful approach to instantaneous timbre control would be some means for automatic programming of the salient parameters. An especially compact medium for the instantaneous representation of a parameter is an analog voltage, which can be conveniently transported from the output of a "parameter synthesizer" to the input of a harmonic tone synthesizer. Thus a voltage-controlled additive harmonic tone synthesizer (which I will refer to from now on as a harmonic tone generator) is a device with 2n + 1 analog voltage inputs--corresponding to the fundamental frequency and the amplitudes and phases of n harmonics--and a single output which is the sum of the n controlled harmonic frequencies. By contrast, a voltage-controlled subtractive synthesizer would have one input for control of the fundamental frequency and several others to control the shape of a formant filter; the shape can be effectively controlled by altering the intensities of the outputs of several fixed band pass filters, or more flexibly by controlling the bandwidths, center frequencies, and intensities of several voltage-variable filters. Hence, in a voltage-controlled parameter system, the information-bearing parameters, which are physically separated from the main synthesizer system, can be generated by whatever means is available and according to whatever designs are desired. The only limitations on the specification of a parameter as a function of time are those imposed by the control response characteristics of the synthesizer and the capability of the parameter voltage generators to produce outputs corresponding to arbitrary functions of time.

Circuitry for generating control parameters is discussed in Chapter 3. It should be noted that certain classes of attack and decay shapes are easy to generate electronically, and that by using these shapes, the effect of independent control of the harmonic envelopes can be conveniently investigated. Also, sound effects such as vibrato, trill, "buzz", and "clangorous sound" can be produced readily by slow and fast sine and square wave frequency modulations.

1.6 Some Applications of Harmonic Tone Synthesis

The input voltages to a harmonic tone synthesizer can be regulated in a number of ways, including direct human operation. For instance, a keyboard can be wired with a set of resistors and a voltage reference in such a way that for each key depressed a voltage is generated which, when applied to the frequency control input, determines a pitch corresponding to the key. Similarly, a conducting wire can be arranged to make contact with a biased wire wound resistor so as to produce both discrete and continuous pitch-voltage changes according to the position of the hand. It would be very difficult for one person to control simultaneously all the harmonic intensities without some aspects of the control being pre-set. One method is to have the voltage generated by the keyboard or conducting wire actuate a set of envelope synthesizers with attack and decay times controlled by means of several stop positions. Another method is touch control, where the pressure exerted on the keyboard produces a proportional voltage for control of the intensities. Other effects, such as vibrato, can be varied by manual manipulation of the frequency and amplitude controls of a sine generator. In this form, the voltage-controlled harmonic synthesizer becomes a very flexible monophonic musical instrument, capable of generating an unusual variety of timbres at the disposal of the musician performing it.

For electronic music, contours of pitch and timbre ought to be more rigidly controlled with respect to time. These contours, in the form of time-varying analog voltages, can be derived from some sort of "synthesizer" of musical parameters, analogous to the paper roll used in a player piano. When a large number of parameters is being generated simultaneously to control a number of single sounds , or, for that matter, more complex homophonic effects, it is of utmost importance that time synchronization be maintained. This condition is satisfied for either an analog parallel tape system such as a clear celluloid with several distinct channels which are darkened proportionally to each parameter intensity or by means of digital-to-analog converters in conjunction with a coded system such as punched paper tape, digital magnetic tape, or direct digital computer output. With this type of synthesizer system in operation, made feasible by the voltage-controlled circuitry, the conventional time-consuming method of adjoining musical events by tape splicing would be less frequently required.

As mentioned in Section 1.1 harmonic synthesizers have useful application for studies in sound perception. Data derived from a thorough investigation of the harmonic tone may shed light on certain properties of the reaction of the ear which have not been heretofore tested, particularly the response of the ear to complex transient signals. Also, through experience in using these devices in electronic music we cannot help but gain insight into the problem of predicting the specification of a sound when given its imagined subjective effect (e.g., in the mind of the composer).

In general, musicians need to acquire an education about the quantifying of timbre comparable to their traditional education about pitch organization. Since their contact with a subject is most direct when sound and physical representations and not mathematics are involved, a harmonic tone synthesizer in conjunction with a good sound playback system and, perhaps a visual spectrum analyzer, would constitute a good educational tool for an explanation of timbre in quantitative terms. People who work in scientific areas may also be interested to know the relation between a mathematical Fourier series and the waveform and sound it produces.

Chapter 2

DESIGN OF A HARMONIC TONE GENERATOR

2.1 Introduction

As discussed in the previous chapter, a harmonic tone generator, which produces a harmonic series of frequencies with fundamental frequency, harmonic amplitudes, and harmonic phases independently adjustable, is essential for the further development of electronic music and science-oriented music research. However, before describing a practical realization of this instrument, it is appropriate to present in detail some of the basic design decisions and factors which restrict the design possibilities. To the designer of a harmonic tone generator the following questions must be posed: (1) What is a convenient and flexible form of the operation of this device? (2) What are some possible alternatives for its realization? (3) What requirements on tolerances and ranges of the defining parameters are necessary because of musical and perceptual considerations?

2.2 Operational Form

For the production of musical sound the operation of an additive synthesizer is in a very convenient form if the parameters which define the output correspond to the important musical parameters: pitch (frequency of the fundamental) and tone quality (distribution of the amplitudes and, to a much lesser extent, the phases of the harmonics). It should also be possible to vary these parameters independently of one another. Furthermore, in order to achieve a most flexible and instantaneous means for controlling these parameters, it is advantageous that they be remotely and independently programmable by external analog DC voltages rather than by means of manual or servo control of variable physical components. Hence, this is a device with several inputs and a single output. A representation of its output voltage can therefore be written explicity in terms of its input voltages:

$$V_{out} = \sum_{k=1}^{n} a_k(t)\cos(2\pi\, k\int f(t)dt + \phi_k(t)) \tag{2-1}$$

$$a_k(t) = c_k V_{a_k}(t), \quad \phi_k(t) = pV_{\phi_k}(t), \quad f(t) = q(V_f(t))$$

a_k is the instantaneous amplitude of the kth harmonic,

V_k is the kth harmonic amplitude control voltage, and

c_k is the constant relating the two (set by an attenuator).

ϕ_k is the instantaneous relative phase shift of the kth harmonic,

V_{ϕ_k} is the kth harmonic phase control voltage,

p is a constant which relates ϕ_k and V_{ϕ_k}.

f is the instantaneous fundamental frequency,

V_f is the fundamental frequency (i.e., pitch) control voltage,

q is a function which relates f and V_f.

n is the number of harmonics which are generated.

Since any voltage source, such as a sine generator, can be used as a fundamental frequency, harmonic amplitude, or phase input, the parameters which define the output normally vary with time.

2.3 Three Possible Methods for Additive Generation of Harmonic Tones

The first method given below was adopted for the construction of a practical harmonic tone generator, which is described in Section 2.5. However, two other methods were considered. These will be described more briefly.

It should be noted that none of these methods utilize physical moving parts. They are completely electronic. Harmonic generators have been built using AC generators driven by harmonically geared motors.[6] Also, a few musical instruments have employed electro-mechanical additive synthesis techniques For example, tones produced by the Hammond Organ are formed by the superposition of six individually controlled partials.[7] These partials are not tuned to strict harmonic ratios, however, since the frequencies are borrowed from the equal-tempered scale. A particularly elaborate harmonic tone synthesizer was constructed by Bell Telephone Laboratories in 1940.[8] The waveforms of 100 harmonic sine waves were recorded on separate channels of a rotating magnetic drum. These sine waves were then read from the drum and filtered to increase their purity. Their individual amplitudes were regulated by means of a set of 100 linear attenuators mounted on a single panel. The fundamental frequency of the harmonic tone generated could be altered by changing the drum's speed of rotation.

a. Hetrodyne Method

The generation of harmonics by the hetrodyne method involves the use of two ultrasonic carrier frequencies: f_c, which is fixed, and f_u, which is variable. f_u is variable between the values f_c and $f_c + f_{max}$, where f_{max} is

---

6. Miessner, B.F.: "Electronic Music and Instruments," Proc. Inst. Rad. Eng., 24(11), 1427-1463, 1936.

7. Dorf, R.H.: Electronic Musical Instruments, Radio Magazines, Inc., Mineola, N.Y., 1958, pp. 25-45.

8. Fletcher, H.: "The Pitch, Loudness and Quality of Musical Tones," Bell Telephone System Monograph No. 1397, 1946.

an audio frequency. The audio beat tone which results from mixing f_c and f_u is $f_a = f_u - f_c$. f_a is the fundamental frequency of this generation scheme, and it varies between 0 and f_{max}. If the kth harmonics of f_u and f_c, namely kf_u and kf_c, are generated and mixed, the frequency of the resulting beat tone is $kf_u - kf_c = k(f_u - f_c) = kf_a$, which varies between 0 and kf_{max}.

This gives a method for deriving a series of audio harmonics from the corresponding harmonic series of two ultrasonic frequencies. The attendant advantage is that the ultrasonic harmonics can be separated by means of filters over a reasonably wide frequency range. There are several methods of performing this separation, but the logically least complicated way is to separate the harmonics generated by a special waveform such as a pulse wave. Since f_c is a fixed frequency, its harmonics can be easily separated by narrow band filters. However, there are restrictions on the response characteristics of the band pass filters used to separate the harmonics of f_u which depend on the values of f_c, f_{max}, and n, the number of harmonics to be generated. Even if perfect filters with rectangular response characteristics were available, the relation,

$$n(f_c + f_{max}) < (n + 1)f_c \qquad (2\text{-}2)$$

must be satisfied to allow the separation of the nth harmonic from the (n+1)st harmonic of f_u over its entire frequency range. Given that n harmonics are to be separated over the entire frequency range of the audio fundamental, Equation (2-2) can be solved to yield the following restriction on the value of the "carrier frequency" f_c:

$$f_c > nf_{max}. \qquad (2\text{-}3)$$

In actual practice, the use of filters of reasonably low complexity dictates that f_c be at least twice nf_{max}.

If kf_u is the drive frequency input to a "chopper-type" phase detector and the detected signal is a sine wave of frequency kf_c, there results a pure audio frequency output, $(V_{a_k}/\pi)\cos(k2\pi f_a + \phi_k)$, where V_{a_k} and ϕ_k are the amplitude and phase of the fixed frequency kf_u, respectively. The phase of kf_c can be controlled by means of a phase modulator circuit, and the amplitude can be controlled by means of a chopper circuit which samples the amplitude control voltage for one-half cycle at the rate of kf_c in order to produce a variable-amplitude square wave. This is converted to a sine wave by a tuned circuit immediately preceding the phase detector. The entire process for generating one audio harmonic with voltage-controlled amplitude and phase is depicted in Figure 2.1.

Pure audio harmonics can also be produced if the phase detector is driven at the rate f_u (instead of kf_u), and care is taken to insure that the detected signal is free from other frequencies harmonically related to f_c. When this method is used, the need for band pass filters to separate the harmonics of f_u is eliminated. The output of the phase detector for each harmonic is maximized if the samples of the detected signal, which occur every $1/f_u$ seconds, are of optimum duration α_k/f_u. Its precise form can be determined through analog multiplication of the signal $V_{a_k}\cos(k2\pi f_c t + \phi_k)$ and the Fourier series representation of the appropriate pulse train:

$$V_{pd\,out} = V_{a_k}\cos(k2\pi f_c t + \phi_k)\left[\frac{1}{2} + \frac{1}{\pi}\sum_{j=1}^{\infty}\frac{\sin(j\pi\alpha_k)}{j}\cos[j2\pi(f_c + f_a)t]\right] \qquad (2\text{-}4)$$

$$= \frac{1}{2\pi} V_{a_k}\frac{\sin(k\pi\alpha_k)}{k}\cos(k2\pi f_a - \phi_k) + \text{high frequency terms.}$$

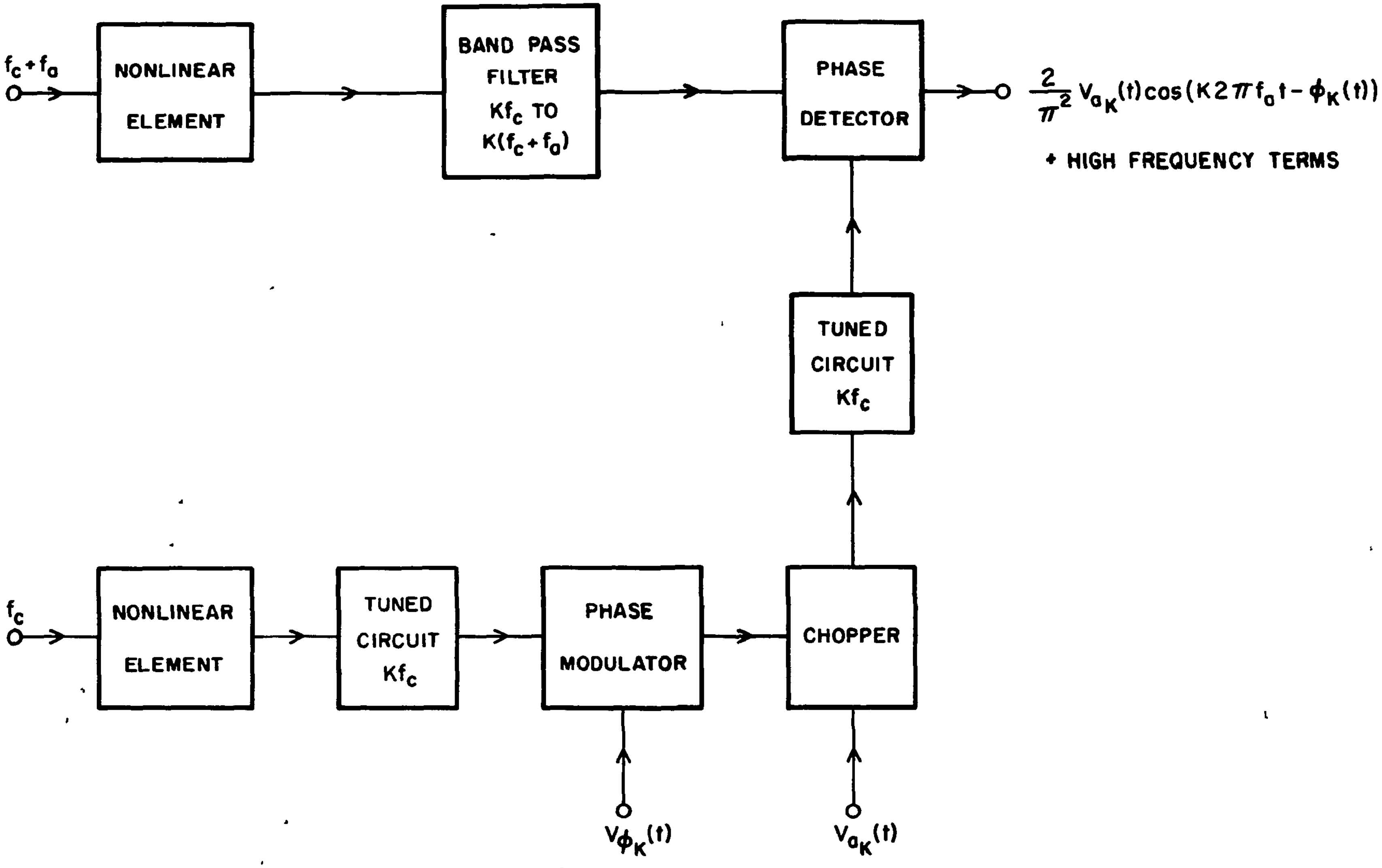


Figure 2.1 Block Diagram for Generator of Audio Harmonic with Voltage-Controlled Amplitude and Phase.

The amplitude of the audio component of the output is maximized when the ratio of the sample duration to the interval of repetition is given by $\alpha = m/2k$, where m is the largest odd integer not exceeding k.

If it is not required that the audio harmonics be generated independently, i. e., if an output consisting of the sum of the modulated harmonics is sufficient, there results a considerable savings in circuitry. In this case, the harmonics of f_c are formed, phase and amplitude modulated, and summed prior to detection by a single phase detector which is driven at the frequency f_u. The audio output is:

$$\frac{1}{2\pi} \sum_{k=1}^{n} V_{a_k} \frac{\sin(k\pi\alpha)}{k} \cos(k2\pi f_a t - \phi_k) \qquad (2\text{-}5)$$

where α/f_u is the phase detector sample time chosen to produce suitable amplitudes for all of the n harmonics. The harmonic amplitudes corresponding to a standard control voltage level should be equalized by a pre-emphasis of the weaker harmonics of f_c before summation. The V_{a_k}'s must be so scaled that the composite signal does not exceed the peak value upon which the phase detector circuit can linearly operate. Since α is not optimum for each harmonic, and each V_{a_k} is reduced, the resulting amplitudes of the audio harmonics are considerably less for a system with a single phase detector than for a system utilizing individual phase detectors for each harmonic (with equivalent circuitry).

In any case, the variable ultrasonic frequency f_u is derived from the output of a voltage-controlled oscillator, so that the fundamental audio frequency is a function of a pitch control voltage.

The chief advantages of the hetrodyne method are: (1) The harmonics are automatically synchronized over the entire range of the fundamental; and (2) each of the important parameters, the harmonic amplitudes and phases and the fundamental frequency, is voltage-controlled. A disadvantage of this method is that the circuitry is quite complex, requiring separate circuitry for each harmonic. Also a practical disadvantage is the difficulty in realizing an ultrasonic v.c.o. with sufficient frequency stability. However, this can be improved if f_u is derived, by single-side-band operation, from the output of a more stable lower frequency v.c.o. In this case, care must be taken to prevent harmonics of the lower frequency from coupling into the ultrasonic system and by hetrodyne action producing inharmonic image frequencies in the audio output.

b. Multiplier-Subtracter Method

A second method for generating harmonics which was also given some consideration involves a multiplier-subtracter process. This method for producing independent harmonics does not employ frequency sensitive elements and has no theoretical frequency range limitations. It utilizes the trigonometric identity,

$$\cos(j\omega t)\cos(k\omega t) = \tfrac{1}{2}\cos(j-k)\omega t + \tfrac{1}{2}\cos(j+k)\omega t, \tag{2-6}$$

which can be abbreviated by

$$jxk \longrightarrow (j-k) + (j+k).$$

The extraction of five harmonics from a fundamental sine tone is then accomplished by the following procedure:

$$1x1 \longrightarrow 0 + 2 \longrightarrow 2,$$

$$1x2 - 1 \longrightarrow 1 + 3 - 1 \longrightarrow 3,$$

$$2\times2 \longrightarrow 0 + 4 \longrightarrow 4,$$

$$2\times3 - 1 \longrightarrow 1 + 5 - 1 \longrightarrow 5, \text{ or } 1\times4 - 3 \longrightarrow 3 + 5 - 3 \longrightarrow 5. \qquad (2\text{-}8)$$

It is clear, therefore, that any number of harmonics can be separated by means of a system consisting solely of multiplier and subtracter circuits. An audio voltage-controlled oscillator is the source of a fundamental sine tone and other multipliers are used to control the individual harmonic amplitudes. The even harmonics can be derived by squaring circuits so that for multipliers operating by means of the "quarter-square" principle (each multiplier containing two squarers), the equivalent of one multiplier can be used to derive two even harmonics.

There are, unfortunately, several obvious drawbacks to this type of harmonic generator: (1) Multipliers of sufficient accuracy and wide frequency range are very expensive. (2) Cancellation of harmonics by subtracting signals is subject to phase and amplitude errors. (3) Errors in producing higher harmonics are bound to accumulate and exceed tolerable limits. (4) The relative phase shifts of the harmonics cannot be easily controlled.

c. Voltage-Controlled Oscillator Method

Finally, as a third possible process a set of identical sine wave v.c.o.'s with either linear or exponential frequency-versus-voltage characteristics might be arranged to track a master control voltage so that the individual output frequencies would be at fixed ratios to one another. In particular, the ratios can be harmonic. Suppose that the response curves are linear, so that for each oscillator $f_{out} = cV_{in}$. Suppose further that each of the v.c.o.'s is preceded by an attenuator. Then, if a master voltage V_m is applied in parallel to the inputs of these attenuators and they are set so that the input voltages to n v.c.o.'s are V_m/n, $2V_m/n$, $3V_m/n$, . . . , $(n-1)V_m/n$, V_m, consecutively, the output frequencies are in the relation 1:2:3: . . . : n-1:n, respectively. Similarly, suppose that the response curves are exponential, so that $f_{out} = k_1e^{k_2V_{in}}$ for each oscillator and that the v.c.o.'s are provided with voltage summers at their inputs. Then, if V_m is applied in parallel to the summer inputs and summed consecutively with the voltages 0, $\ln(2)/k_2$, $\ln(3)/k_2$. . . , $\ln(n)/k_2$, the v.c.o output frequencies are again in harmonic relation.

Three advantages of either of these systems are: (1) All sub-units are identical; (2) nonlinear networks for extracting harmonics are not required, and (3) the oscillators may be used for applications other than harmonic synthesis (for generating inharmonic tones or separate tones). On the other hand, the requirements with regard to frequency stability and frequency tracking over a wide range are very stringent. A 40 to 20,000 cps frequency range is useful, but very small frequency errors (2.5 cps or .2 percent, whichever is greater)[9] are perceptible Moreover, the oscillators are not synchronized (although synchronization apparatus could be introduced for applications which require a limited frequency range), and so the relative phases of the harmonics are continually changing with respect to one another According to more than one researcher,[10,11,12] phase changes in the higher numbered harmonics do affect the perception of tone quality, and it follows that the timbre of a tone with non-synchronized harmonic partials is not necessarily a fixed quantity with respect to time.

---

[9] Shower, E.G and Biddulph, R : "Differential Pitch Sensitivity of the Ear." J. Acoust Soc. Am , 3(2), 275, 1931.

[10] Mathes, R.C. and Miller, R.L. : "Phase Effects in Monaural Perception," J. Acoust. Soc Am , 19(5), 780-797, 1947.

[11] Licklider, J C. : "Effects of Changes in the Phase Pattern Upon the Sound of a 16-Harmonic Tone," J. Acoust. Soc. Am., 29(6), 780 (abstract), 1957

[12] Schroeder, M R. : "New Results Concerning Monaural Phase Sensitivity," J Acoust Soc. Am , 31(11), 1579(abstract), 1959.

If the v.c.o.'s are "beat-frequency" oscillators, there is a particularly convenient method for amplitude modulating the individual harmonics (as described in Section 2.3a), otherwise, the amplitudes of the individual harmonics must be controlled at the outputs of the v.c.o.'s which generate them.

2.4 Some Design Considerations

A precise definition of the transfer characteristics of a harmonic tone generator, according to Equation (2-1), requires that the specifications be given for the values of n, p, and q, for the range of values for f(t), $V_f(t)$, the $V_{a_k}(t)$, c_k, and $\phi_k(t)$, and for the tolerances permitted these values as well as other criteria relating to the "closeness" of Equation (2-1) as a description of the measured output.

a. Number of Harmonics

The question arises as to how many harmonics are needed to create aurally significant tones and how many harmonics should be included in an additive synthesis system before a point of "audible diminishing returns" is reached. There are some data available on conventional musical instrument tones which may help answer this question. The second column from the left in Table I. gives the lowest fundamental frequency normally played on each instrument,[13] the next column gives the frequency at which the formant response of the instrument is 30 db below the peak resonant value,[2] the last column on the right gives

---

[13] Culver, C.A. . *Musical Acoustics*, McGraw-Hill, New York, 1956, p. 95.

Instrument	f_1 (cps)	f_2 (cps)	$^{f_2}/f_1 = n$
Flute	261	3000	11.5
Oboe	233	8000	34.3
English Horn	165	4000	24.3
Bassoon	58	2000	34.5
Trumpet	165	3000	18.2
Trombone	82	2000	24.4
Tuba	41	1200	29.3
French Horn	62	800	12.9
Violin	196	7000	35.7
Viola	131	3000	22.9
Cello	65	800	12.3
String Bass	41	1500	36.6

Table I

Estimate of Number of Harmonics Required to Simulate the Tones of Several Musical Instruments Over Their Complete Ranges (n). f_1 = lowest fundamental frequency. f_2 = highest frequency at which formant response is less than 30 db below peak formant response.

the quotient of these two values, i.e., an estimated measure of the maximum number of harmonics required to simulate each instrument's steady-state tone. Note that according to this estimate, the flute, cello, and French horn can be simulted over their complete ranges with twelve or thirteen harmonics whereas many of the other instruments require over thirty harmonics.[14] These figures may be unduly pessimistic due to masking factors, and subjective experience with additive synthesis systems may be the only means of settling this question. I feel that twelve harmonics are probably sufficient for initial investigations into the perceptive properties of harmonic tones. Although the six harmonics which are now producible on the Harmonic Tone Generator model described in Section 2.5 are inadequate for creating a complete range of steady-state timbres, this generator is still interesting to many composers of electronic music because of its unusual flexibility for creating changes in the frequency and amplitude parameters which conventional instruments cannot simulate.

b. Control of the Fundamental Frequency

Four factors to consider when specifying the control of fundamental frequency are: frequency range, frequency stability, the frequency-vs.-voltage control curve, and the time lag between frequency changes and discrete control voltage changes

For any practical voltage-controlled oscillator which is to be the source of a variable frequency, instability generally increases with increasing frequency range; hence, the frequency range must be limited. It is evident that taste and perceptual limits have influenced the choice of the upper pitch limits for ordinary musical instruments. Of the conventional instruments, only the piano and pipe organ generate fundamental frequencies above 4000 cps., and only the violin, harp, flute, piccolo, and xylophone exceed 2000 cps.[15] Also, listening tests indicate that the perception of musical pitch falls off rapidly above 4000 cps.[16] If the maximum fundamental (first harmonic) frequency output producible by a harmonic synthesizer is chosen to be f_{max} = 2000 cps., a 4000 cps. range is easily accommodated through setting the odd harmonics to zero amplitude.

---

[14] V.O. Knudsen proposed in 1936 that a generator of 40 harmonics be produced See his "Modern Acoustics and Culture," University of California Press, Berkeley, California, 1937, pp. 22-31.

As mentioned above, pitch discrimination tests made under optimum listening conditions indicate that a suitable criterion for frequency stability is: 2.5 cps. or .2 per cent whichever is greater.[9] This should be maintained over a range of room temperatures likely to be encountered in a temperate climate, i.e., 60 to 90°F. The maximum short-term frequency deviation which is tolerable should also be specified, and in this case tape recorder flutter data are quite applicable. An acceptable criterion for frequency "jitter" seems to be: .3 per cent r.m.s. integrated over components between 1 and 200 cps.[17]

---

[15] Olson, H.F. : Acoustical Engineering, D. Van Nostrand Co., New York, 1957, p. 610.

[16] Bachem A. : "Chroma Fixation at the Ends of the Musical Scale," J. Acoust. Soc. Am., 20(5), 704-705, 1958.

[17] Haynes, N.M. : Elements of Magnetic Tape Recording, Prentice-Hall, Englewood Cliffs, N.J., 1957, pp. 183-193.

pitches has a great deal of validity for musicians,[18] despite the fact that the number of just-detectible frequency changes per octave increase significantly from the low to high frequency ranges.[19] Our Western equal-tempered scale is, of course, strictly logarithmic. Therefore, an anti-log or exponential curve describing the frequency as a function of control voltage would be a logical type to specify.

From the standpoint of music system design, both the exponential and linear functions are convenient because of the attendant possibilities for frequency transposition. In the case of exponential control, any sequence of voltages, converting to a pattern of pitches, can be arithmetically added to a fixed-voltage in order to produce a composite pitch control voltage. This in turn, produces an interval transformation of the original pitch pattern proportional to the "augmenting" voltage. Similarly, if this "augmenting" voltage, which is added to the "main" pitch control signal, alternates in such a way as to produce a vibrato or trill with a certain frequency deviation with respect to a given "main" frequency, e.g., a semi-tone, this frequency interval deviation will remain fixed for any "main" frequency. The electronic design is simplified if the pitch control voltage has a single polarity. Therefore, a satisfactory frequency control function would be exponential from 2093 cps. (C_7) to 32.7 cps.(C_1) (about the lowest useful fundamental) and trace some other curve to the point zero volts, zero frequency. In this way, the generator output can be "turned off" by setting the pitch control voltage to zero, independent of the amplitude control.

---

[18] Ward, W.D. : "Subjective Musical Pitch," J. Acoust. Soc. Am., 26(3), 370-380, 1954.

[19] Lewis , D. : University of Iowa Studies in the Psychology of Music, Vol. IV, 1937, University of Iowa Publication, Iowa City, p. 346.

In the case of linear frequency control, transposition of pitch sequences can be implemented by means of attenuation or amplification of the control voltage, rather than by addition of a fixed voltage. However, the superposition of a constant interval trill or vibrato cannot be implemented by a fixed attenuator. The principal advantage of linear control lies in its relative insensitivity to control voltage inaccuracy

When a sequence of discrete pitch control voltages is applied to the harmonic tone generator frequency input, one should not be able to distinguish the presence of a frequency glide between the resulting pitches. The minimum time which gives this impression when a pitch changes from one discrete value to another has been informally measured in our studio to be between 10 and 20 milliseconds.

c. Harmonic Amplitude Control

For simplicity in system planning and in the electronics design, a linear relation between the harmonic amplitudes and their control voltages seems most logical. Although a case can be made for exponential control, since the perception of intensity is close to being a logarithmic response, the usefulness of linear control in producing easily predictable sideband modulations from alternating external voltages should be noted. The maximum modulation rate should be sufficient to allow the production of sharp attacks. The shortest rise time which can be discriminated from the immediate turn-on of a tone is about 1.5 milliseconds.[20] Therefore, a 1 kc modulation bandwidth is sufficient.

The linearity of output intensity of a single harmonic with respect to its control voltage is not critical since no one envelope function takes a priori precedence over another. Nevertheless, for good correlation between the external envelope control and the aural perception of an attack envelope resulting from this control, the linearity should be within 1 per cent full scale. A convenient full scale voltage is 10 volts.

Two other factors which degrade amplitude control are (a) any finite amount of a harmonic output when its control voltage is zero (zero leak); (b) any amount of the control signal which appears at the output (control leak). The zero leak should be sufficiently small compared to the maximum harmonic output as to be inaudible. This requirement for inaudibility depends on the harmonic frequency, normal maximum listening level, and environmental background noise. The Fletcher-Munson curves[21] demonstrate that for a 3000 cps. sine tone heard at a maximum 90 phons in a noise free chamber, the zero leak must be about 90 db down from this level to be inaudible to acute ears, whereas at 50 cps., only a 35 db difference is needed. These curves were derived under ideal conditions, however, and in a room with reasonable background noise, such as the University of Illinois Experimental Music Studio, a 70 db reduction seems to be adequate at any frequency. The amount of amplitude control signal which appears at the output for any modulation frequency should be balanced out to at least 40 db below the output level. One reason for this balance requirement is that during the transient period a sharp attack envelope voltage appearing at the output constitutes a very wide spectrum and will be perceived as an undesirable "click". The click should be low enough in intensity to be masked out by the sound itself.

---

[20] Mathews, M.V. : "The Digital Computer as a Musical Instrument," Science, 142(3592), 553-557, 1963.

[21] Fletcher, H. and Munson, W.A. : 'Loudness, Its Definition, Measurement and Calculation," J. Acoust. Soc. Am., 5(2), 82-108, 1933.

d. Harmonic Relative Phase Control

Statements about the dependence of tone quality on static relative phases are still very speculative, although it now seems certain that there is some effect.[10,11,12] For at least two reasons, it is believed that phase distributions have a secondary effect compared to frequency distributions. First, the sounds of electronic sources and of the steady-state tones of musical instruments are not appreciably altered by playing these sounds through high quality reproducing systems (including loudspeakers) which have widely different phase shift characteristics. Second, standing-wave distributions in small rooms introduce large phase shift variations throughout with the result that the perceived phase distribution of a steady signal should depend on the listener's position in the room and the position of his head. However, it is not common experience for steady state timbres to change significantly as one walks about a small room. In order that this problem may be investigated in greater detail each harmonic should have provision for changing its phase a full 360° as a function of its phase control voltage. On the other hand, it is felt that the installation of the fundamental frequency and harmonic amplitude controls is much more important for music applications and should be completed first.

The effects of sudden phase changes during the attack transient and slow random variations during the steady state are even more uncertain. It is believed that periodic phase variations exceeding 100 cps are not required for this investigation because these variations are perceived in terms of side-band spectral components and destroy the effect of a harmonic tone.

e. Purity of the Output Signal

Besides the requirement that the various parameters contained in Equation (2-1) be represented within a suitable tolerance, there is the requirement that the level of added unwanted frequency components be below a certain tolerable level. These components can be categorized into three types: (1) harmonic distortion of the individual harmonics; (2) discrete inharmonic components (the relative intensity and frequency of these may well depend on the fundamental frequency and the particular harmonic analyzed); (3) additive white noise. A harmonic distortion level of up to 2 per cent (mostly second and third harmonic distortion) is usually not noticeable since subjective harmonics are usually generated within the ear anyway.[22] But inharmonic tones, even when reduced to an intensity of 40 db or more below the harmonic output can be quite disturbingly evident, particularly when these components are far removed in frequency from the harmonics and hence are not sufficiently masked by them. (A possible source of inharmonic components is noted in Section 2.3a). A suitable requirement is that the inharmonic frequencies be 70 db below the harmonic output at all levels and that the additive noise be 70 db below full harmonic output.

---

[22] Hunter, J.L. : Acoustics, Prentice-Hall, Englewood Cliffs, N.J., 1957, pp. 253-256.

2.5 Description of Circuitry for A Practical Additive Sythesis Device, the Harmonic Tone Generator

A solid-state version of the Harmonic Tone Generator, which generates six independently controlled harmonics, was designed and constructed at the University of Illinois Experimental Music Studio according to the "hetrodyne method" described above. The fundamental frequency, the harmonic amplitudes, and the phase shift of the second harmonic are voltage-controlled parameters. A complete but simplified block diagram illustrating the electronic organization of this device is given in Figure 2.2.

Figure 2.3 shows the schematic of the "frequency synthesizer" sub-unit which generates a pair of ultrasonic frequencies: on one channel a fixed 50 kc "carrier" tone and on the other channel a tone whose frequency varies from 50 to 52 kc as a function of the pitch control voltage. This variable frequency is derived from the output of a voltage-controlled oscillator, the frequency of which varies from 7 to 5 kc as the pitch control voltage changes from 0 to -11 volts. The v.c.o. output is mixed nonlinearly with a fixed 57 kc reference tone, and the resulting 50 to 52 kc difference frequency is filtered out. In this way, a relatively high, voltage-controlled frequency is produced with the stability inherent in a lower frequency oscillator.

The circuit diagram for this voltage-controlled oscillator is included in Figure 2.3. The three varicap diodes shown in the circuit act together as a pure capacitance, the value of which varies with the applied voltage, in this

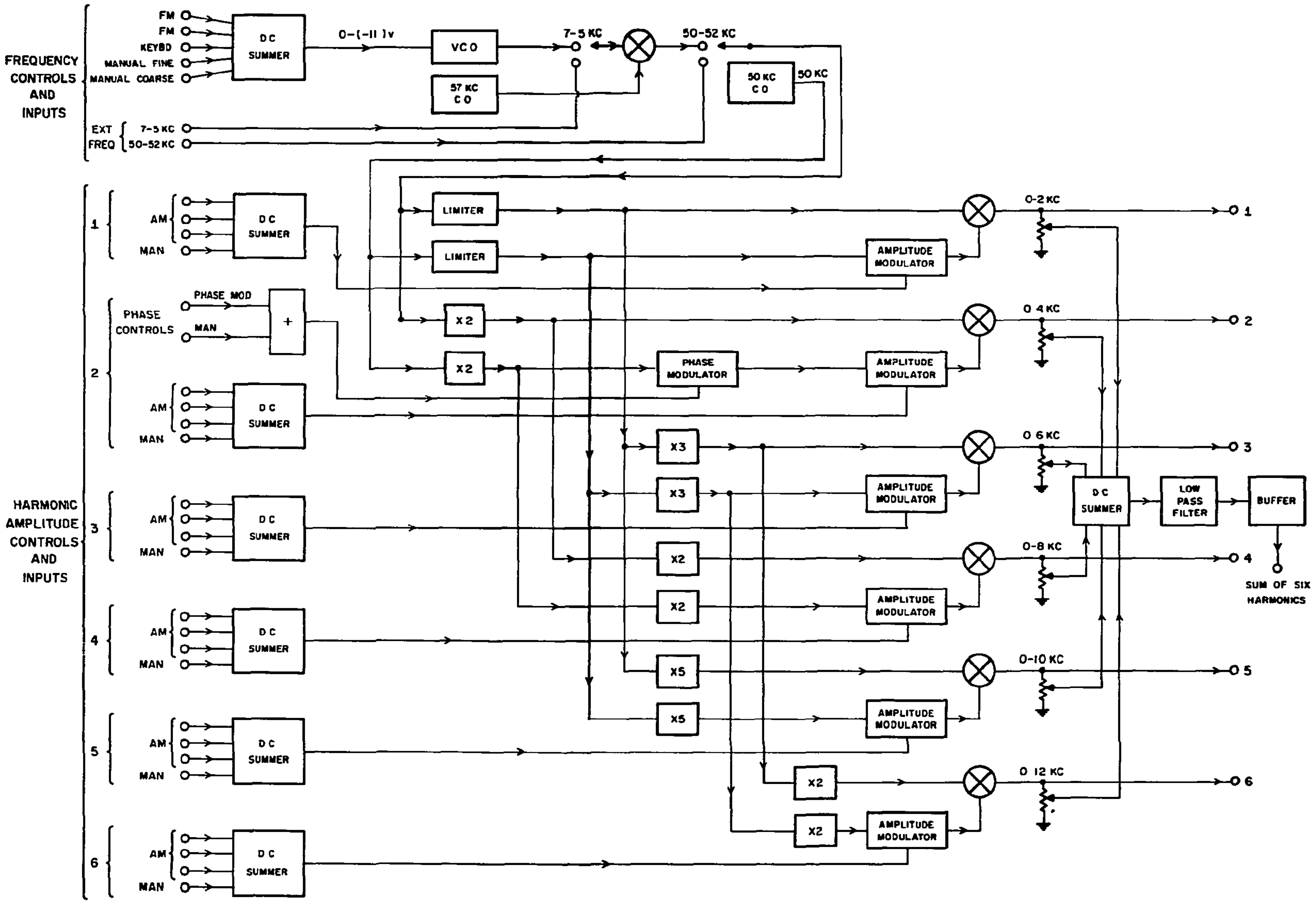


Figure 2.2 Harmonic Tone Generator Simplified Block Diagram.

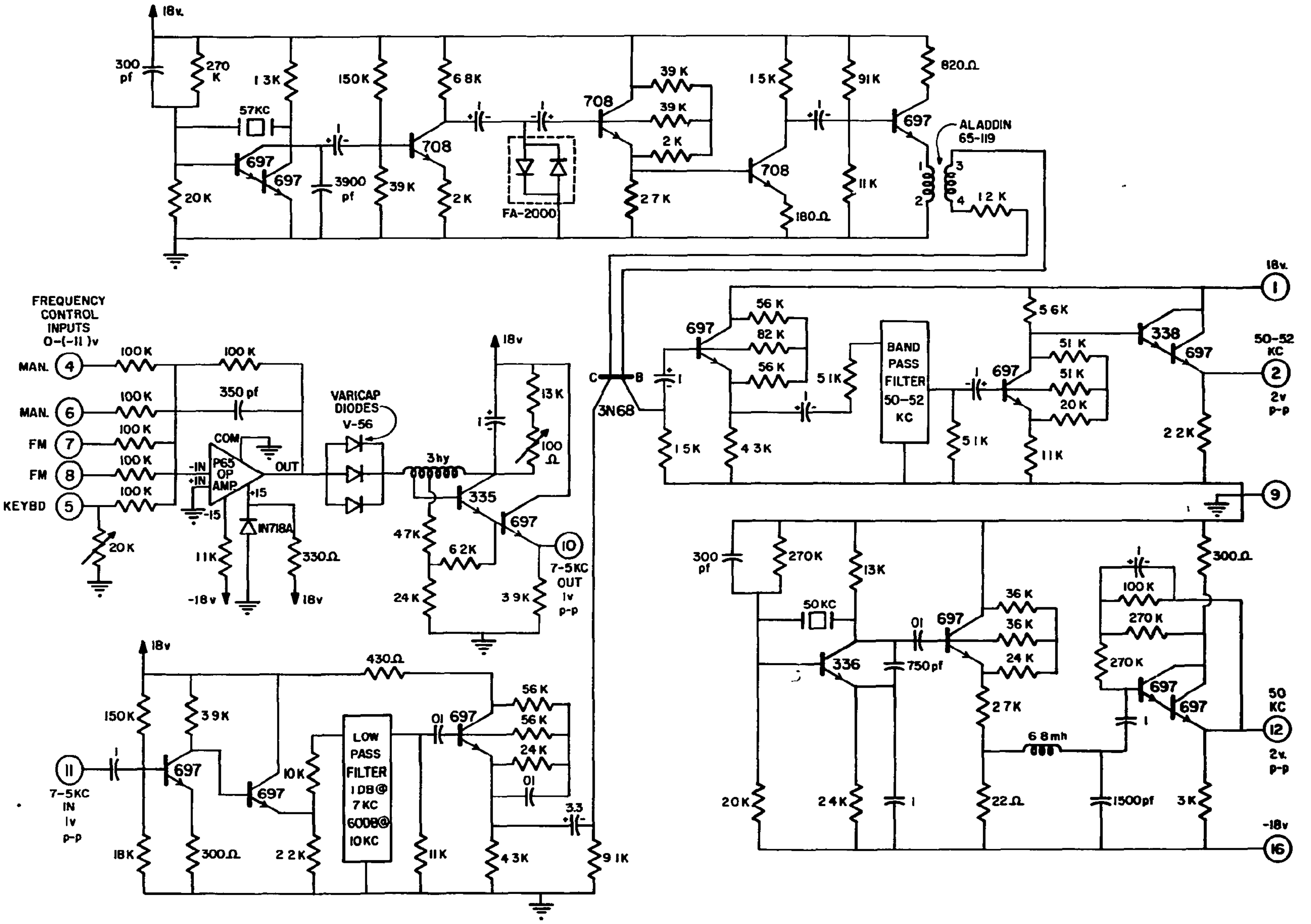


Figure 2.3 Frequency Synthesizer Schematic.

case 11.5 - V_f volts, according to a regularly defined function,[23]

$$C = 3K(11.5 - V_f + V_o)^{-n} , \tag{2-9}$$

where K, V_o, and n are parameters peculiar to a single varicap diode. In addition to this capacitance, the frequency of the oscillator is a function of the inductance and the capacitance of the coil and the transistor input circuit, C_d. It is given by the well known formula, $f_{vco} = \frac{1}{2\pi\sqrt{L(C+C_d)}}$. Therefore, the fundamental frequency of the Harmonic Tone Generator output is given by

$$f_a \doteq 7000 - f_{vco} = 7000 - \frac{1}{2\pi\sqrt{L(C+C_d)}} \tag{2-10}$$

$$\doteq 7000 - \frac{1}{2\pi\sqrt{L(3K(11.5-V_f\ V_o)^{-n} + C_d)}}$$

An empirically derived plot of f_a versus the pitch control voltage V_f is shown in Figure 2.4. It happens that for control voltages between 0 and -4 volts the curve is quite linear, i.e., $f_a \doteq 100\,|V_f|$, whereas between -5 and -11 volts, the frequency is, to a good approximation, an exponential function of voltage, i.e., $f_a \doteq 180\cdot 10^{.1|V_f|}$. This curve could be corrected by means of a diode function generator circuit, to achieve strictly linear or exponential control.

The odd harmonics of the fixed and variable ultrasonic frequencies generated by the "frequency synthesizer" are extracted from square wave distortions of these tones. A symmetrical square wave is obtained by means of a limiter circuit, shown in Figure 2.5, which utilizes matched silicon diodes. Hence, filtering of the third and fifth harmonics is simplified since the output waveform is quite free of even harmonics. Nevertheless, because the relative

---

[23] Weldon, L.A. and Rapski, R.L. : Boost for Electronic Tuning: Part 1," Electronics, Vol. 37, April 6, 1964, pp. 49-53.

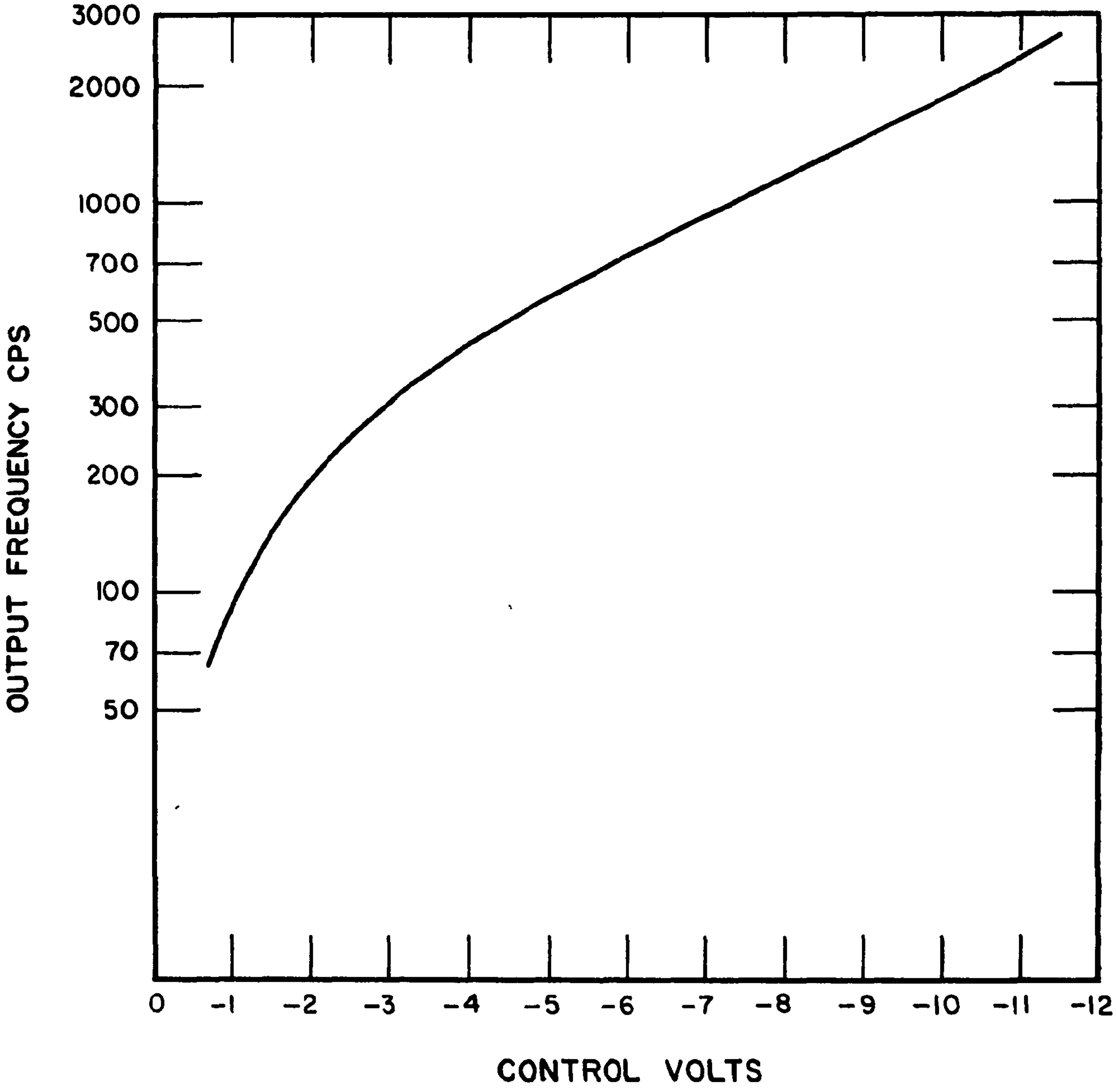


Figure 2.4. Fundamental Output Frequency vs. Input Frequency Control Voltage, Harmonic Tone Generator

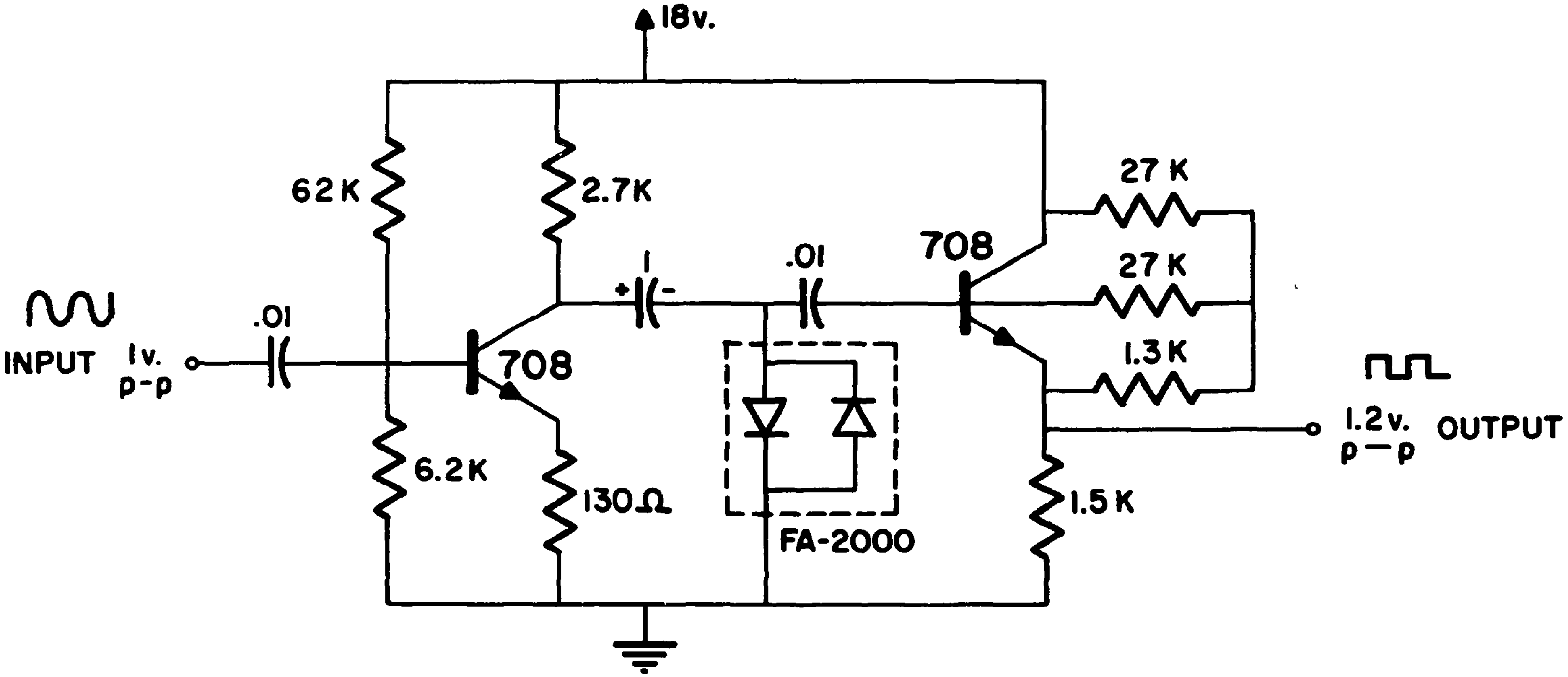


Figure 2.5 Limiter Schematic.

levels of the first, third, fifth, and seventh harmonics, are 0, -10, -14 and -17, respectively, considerable filtering is required to reject unwanted components to a level 60 db below the desired components. By comparison, it is easy to extract the even harmonics. A full-wave rectifier circuit, shown in Figure 2.6, is used as a frequency doubler, in conjunction with a band pass filter. The second harmonic, which is the component desired, and the fourth harmonic, which is to be rejected, have relative levels of 0 and -14 db, respectively, and without adjustment, the undesired odd harmonics are at a level of -30 db or lower.

The harmonics of the fixed "carrier" tone, once separated, are (in general) next subjected to phase and amplitude modulation before they are mixed with the corresponding harmonics of the variable tone to form audio harmonics.

In the construction of the Harmonic Tone Generator, first priority was given to the installation of the harmonic amplitude and fundamental frequency control functions Because the harmonic phases have a much smaller influence on timbre, the installation of phase controls was given lower priority. Consequently, at the present time phase control circuitry has only been completed for the second harmonic. However, phase controls for the other harmonics may be added at a later time.

Figure 2.7 gives the schematic for a phase modulator circuit which is used for the second harmonic. The complex voltage gain of this circuit is

$$\frac{V_{out}}{V_{in}} = \frac{R - j(\omega L - 1/\omega C)}{R + j(\omega L - 1/\omega C)} = e^{j\phi}. \qquad (2\text{-}11)$$

The amplitude of the gain is unity, independent of any of the variables, whereas the phase is given by

$$\phi = -2\tan^{-1}\left[\frac{\omega L - 1/\omega C}{R}\right] \qquad (2\text{-}12)$$

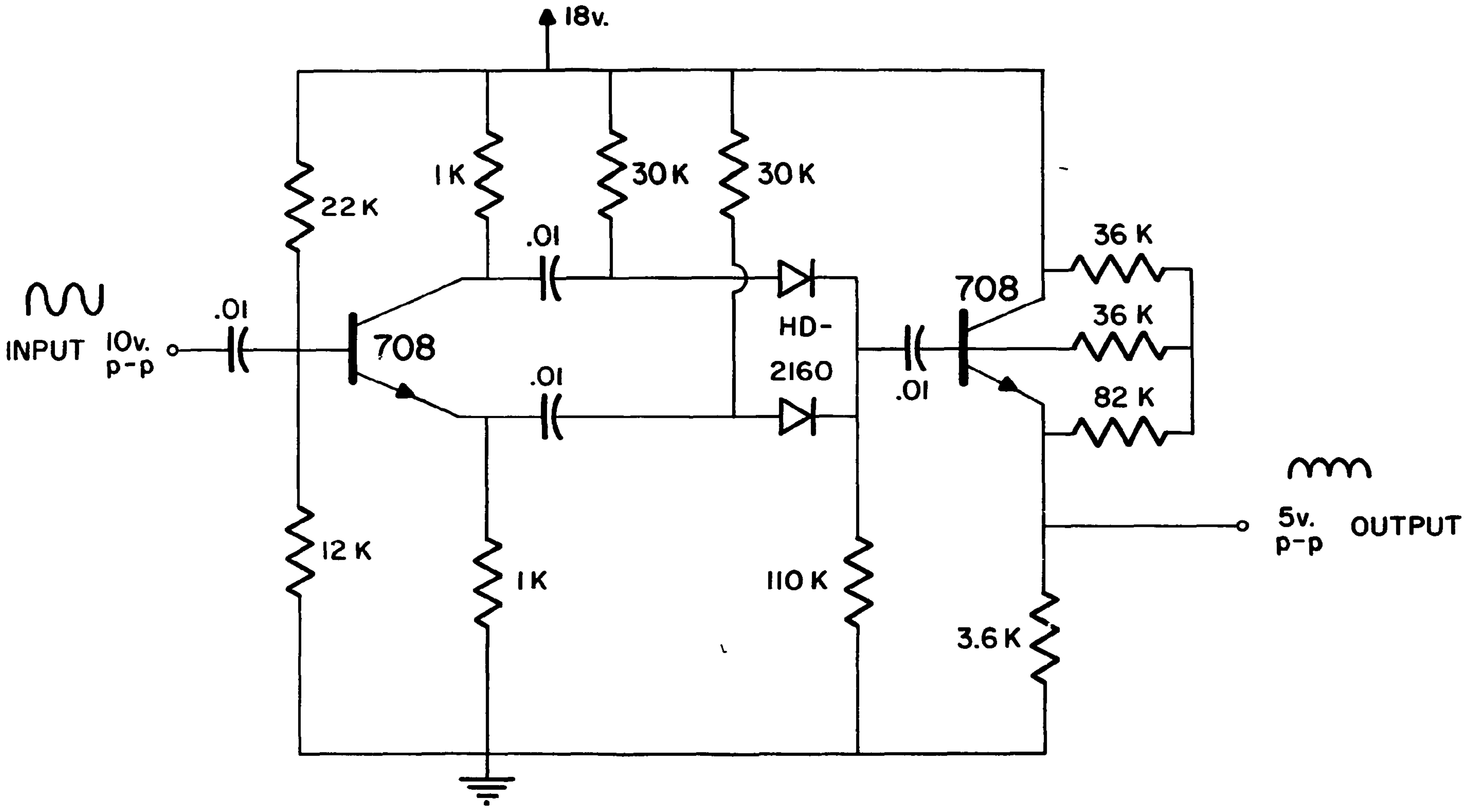


Figure 2.6 Full-Wave Rectifier Schematic

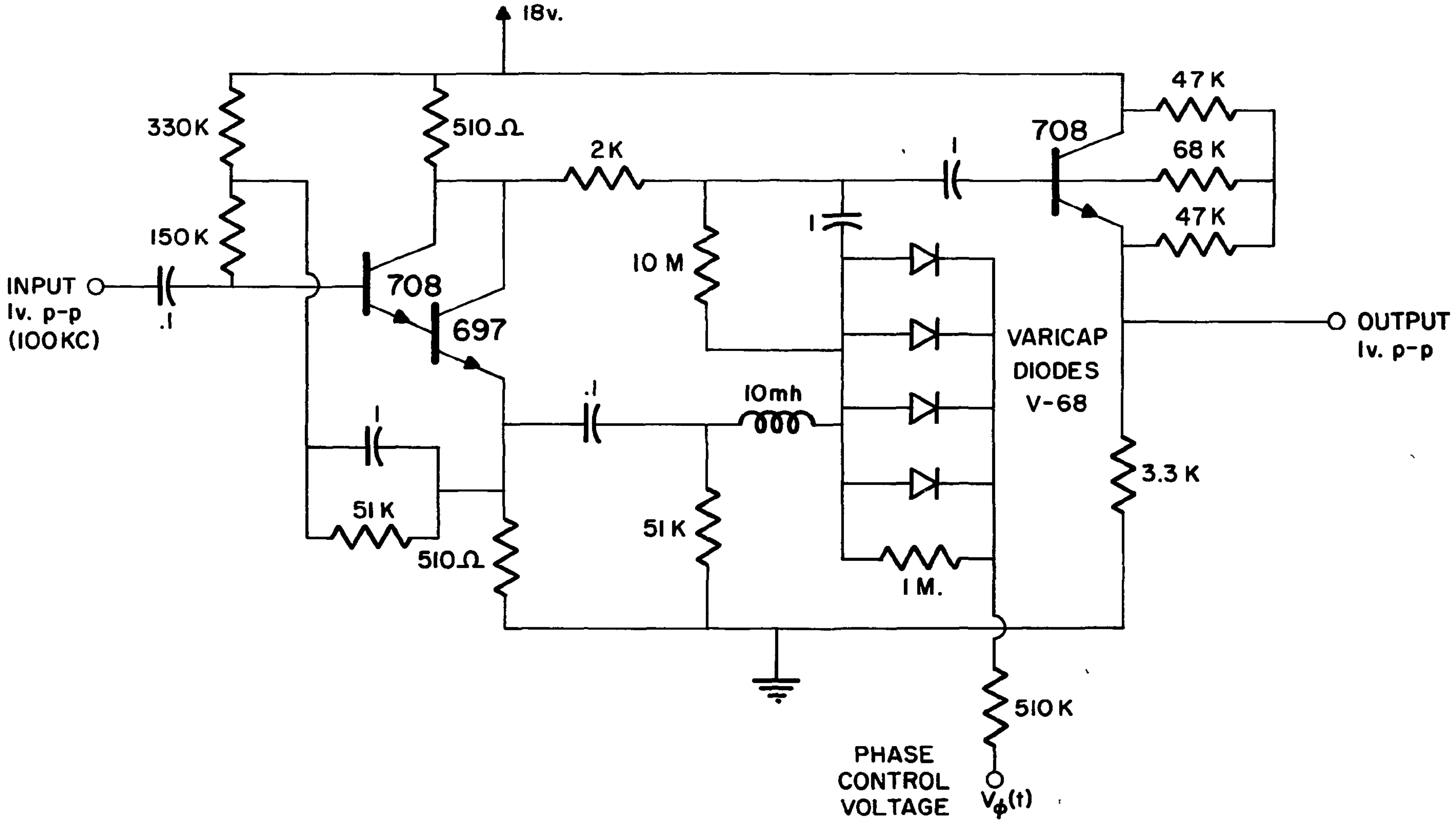


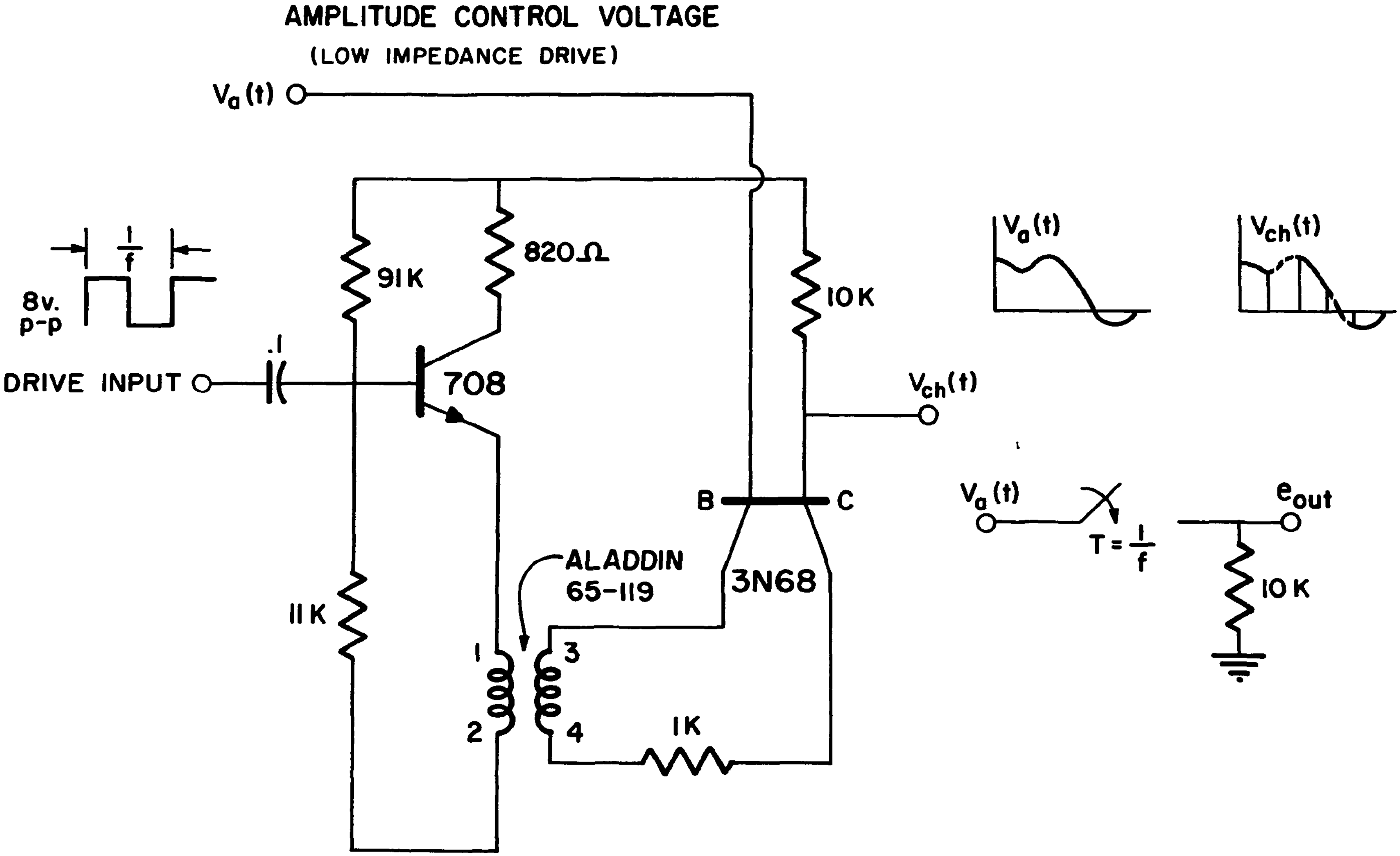
Figure 2.7 Phase Modulator Schematic

Since C is the capacitance of a set of varicap diodes and is a function of the applied voltage, the phase, in turn, is a function of this single-polarity phase-control voltage. The phase shift varies over 180° as the control voltage changes from 0 to -16 volts.

Linear amplitude modulation of each harmonic of the fixed 50 kc "carrier" tone by a control voltage of either polarity is achieved with a "chopper" circuit, which is also useful as a phase detector and as a balanced modulator (used in the frequency synthesizer sub-unit). The diagram for this circuit, which uses a wide-band transformer and a 3N68 integrated chopper, is shown in Figure 2.8 along with its equivalent circuit, a single switch which opens and closes at the rate of the input "carrier" frequency. The output of the chopper alternates between the input voltage $V_{a_k}(t)$ and zero volts at this rate. It follows that the signal produced is identical to that which would result if $V_{a_k}(t)$ were multiplied in analog fashion by a square wave alternating between zero and one. By using the Fourier series representation of this square wave, the output of the chopper can be expressed in terms of its frequency components:

$$V_{ch_k}(t) = V_{a_k}(t)\left(\frac{1}{2} + \frac{2}{\pi}\sum_{i=0}^{\infty}(-1)^i\frac{\cos[(2i+1)(k2\pi f_c t + \phi_k)]}{2i+1}\right) \qquad (2\text{-}13)$$

A tuned circuit which selects the component, $(2/\pi)V_{a_k}(t)\cos(k2\pi f_c t + \phi_k)$, prior to phase detection, completes the amplitude modulator. The tuned circuit Q should be large enough for sufficient rejection of other harmonic components but not so large that a 2 kc amplitude modulation bandwidth cannot be accommodated. When this amplitude-modulated sine wave is applied to the same chopper circuit



ACTUAL CIRCUIT

SWITCH EQUIVALENT

Figure 2.8 Chopper Circuit.

used as a phase detector, the output is

$$V_{pd_k} = \frac{2}{\pi} V_{a_k}(t)\cos(k2\pi f_c t + \phi_k)\left(\frac{1}{2} - \frac{2}{\pi} \sum_{i=0}^{\infty} (-1)^i \frac{\cos[(2i+1)k2\pi(f_c+f_a)t]}{2i+1}\right)$$

$$= (2/\pi^2)V_{a_k}(t)\cos(k2\pi f_a t - \phi_k) + \text{high frequency terms.} \qquad (2\text{-}14)$$

The first term of the last expression represents, then, the only audible component--the kth harmonic with voltage-controlled amplitude and, in the case of the second harmonic, the phase.

The entire process-frequency multiplication, phase modulation, amplitude modulation, and phase detection--is summarized as a block diagram in Figure 2.1.

The relative amount of each phase detector output is set by a manual attenuator before summation of all six harmonic signals is accomplished. A sharp cutoff low pass filter is used to remove the ultrasonic frequencies from the signal. Therefore, the complex audio waveform corresponding to the attenuator and control voltage settings can be observed on an oscilloscope screen.

The pitch and harmonic amplitude control voltages are actually 'composite' voltages, that is to say, each control voltage is the sum of several individual control voltages. For each of the parameters, one of these voltages is adjusted by a manual potentiometer mounted on the front panel. Therefore, the Harmonic Tone Generator can be used as a manually operated generator, as well as a converter of signals from one form into another Also, for both the fundamental frequency and the amplitude controls, it is useful to have the capability for superimposing on 'main' parameter controls signal perturbations which effectively produce vibrato, trill, tremolo, and the like

The front panel layout of the Harmonic Tone Generator (Figure 2.9) is

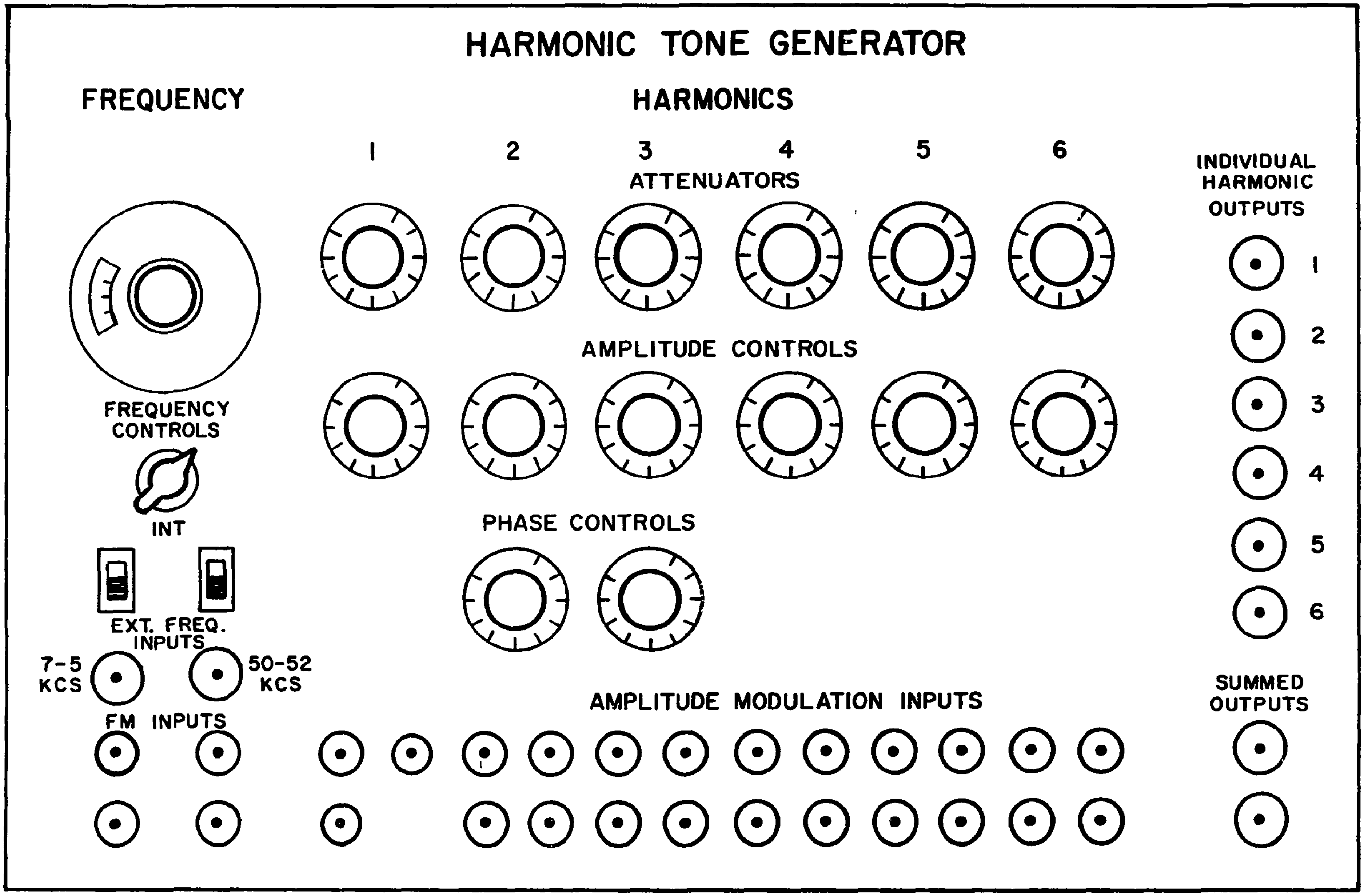


Figure 2.9 Harmonic Tone Generator Front Panel

divided logically into seven columns--one each for frequency, first harmonic, second harmonic, etc. Knobs for manual control of the parameters and for attenuation of the harmonic outputs, and connectors for parameter control inputs, individual harmonic outputs, and summed outputs are mounted on the panel.

In summary, the specifications for the Harmonic Tone Generator, which have been verified by measurement, are:

1. Number of Harmonics: 6
2. Input impedances: 100 K or greater on all inputs.
3. Input voltage ranges: (a) -10 to + 10 volts on amplitude inputs.
    (b) 0 to -11 volts on frequency input.
    (c) 0 to -16 volts on second harmonic phase input.
4. Modulation frequency range: 3KC min. on all inputs except second harmonic phase which is 200 cps.
5. Output range: (a) Fundamental frequency --0 to 2400 cps.
    (b) Harmonic amplitudes --.45 to 250 millivolts rms.
    (c) Second Harmonic Phase -- 180°.
6. Linearity of amplitude control: within 2 per cent max.
7. Total distortion of individual harmonics: less than 2 per cent.
8. Amplitude "zero leak": 55 db below full output.
9. Frequency stability: 2 cps. or .5 per cent whichever is greater for constant temperature (72°F).

10. Output impedance: 30 or 600 ohms.

For many technical reasons, these specifications do not live up to the requirements outlined earlier in this chapter. However, there are no theoretical limitations and we expect to improve these characteristics. Increasing the number of harmonics and installing phase controls for all of the harmonics is primarily a scale-up problem. On the other hand, technical refinement is needed to improve the frequency stability, frequency vs control voltage curve, and amplitude zero leak.

Chapter 3

CIRCUITRY FOR THE GENERATION AND CONTROL OF PARAMETER VOLTAGES

3.1 Introduction

The electrical devices which are briefly described in Section 1.1, items (2) through (6), may be conveniently discussed as pertinent to the control of musical parameters. These devices were constructed primarily for the purpose of generating DC voltages to be used for the time-envelope control of the harmonic amplitudes and fundamental frequency of the Harmonic Tone Generator. However, as is mentioned in Sections 3.5c and 4.1, they may be used in conjunction with other devices which have been built or acquired by the Experimental Music Studio. Figure 3.1 shows a simplified block diagram of the parameter control system and the possibilities for switching between the various units, including the Time-Envelope Networks, the Time-Delay Gate Networks, the keyboard, and the AC-DC converters. The analog multiplier, though it is used for parameter control, is not shown because it is functionally independent of the other units and not attached to the network of switches.

It is possible to build a device which will generate over a finite interval of time a more or less arbitrary DC signal. Two methods for accomplishing this are the photo-waveform method, by means of which an envelope signal can be produced corresponding to a wave form traced upon an oscilloscope screen, and the universal function simulator method, whereby the envelope signal is specified (by adjustment of appropriate controls) at a number of discrete points within the finite time interval. Although from the scientific investigator's

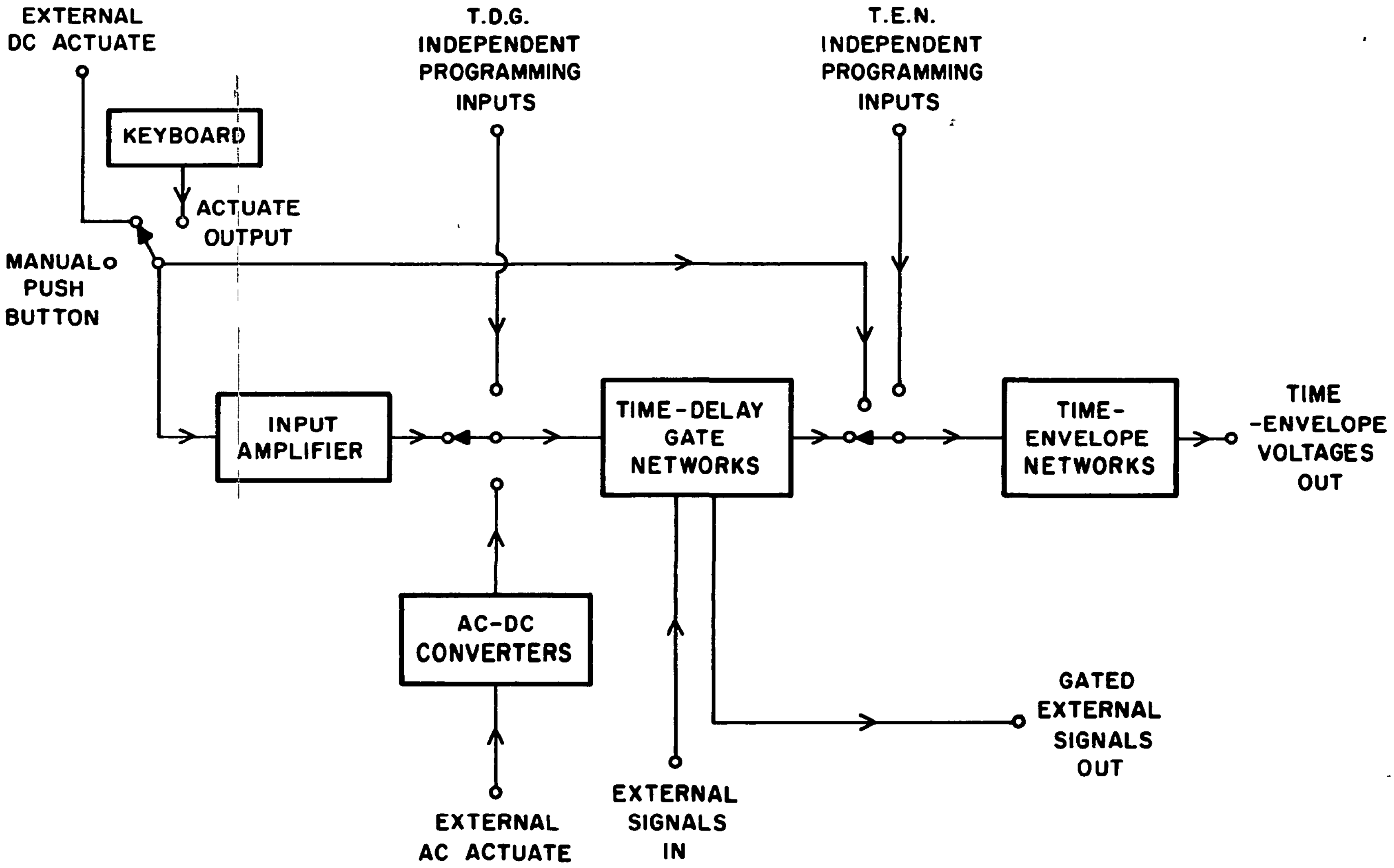


Figure 3.1 Parameter-Control Voltage Synthesizer Simplified Block Diagram.

point of view the ability to specify arbitrary parameter contours is highly desirable, there are two principal reasons why I decided, for the present, not to use these types of units in favor of a less elaborate system for control voltage synthesis. First, since in general several distinct parameter voltages are required as inputs to the Harmonic Tone Generator--and even more will be required if this instrument is to be expanded in the future--, the cost of a set of arbitrary function generators, which are individually expensive would be high. Second, the setting-up of several arbitrary parameter contours using the methods described would be both time-consuming and clumsy. In electronic music, these are important considerations. Consequently, I decided to build a set of modules which would allow the control of the following "sub-parameters" which are believed to have perceptual significance for music:

(1) τ_1, the delay time, i.e., the time of beginning of the envelope signal relative to an arbitrarily placed time origin.

(2) τ_2, the attack time, i.e., the duration of the attack transient.

(3) τ_3, the decay time, i.e., the duration of the decay transient.

(4) A, the relative amplitude of the envelope signal.

(5) p(t), an oscillatory perturbation signal whose amplitude and frequency are described by their own sets of sub-parameters. (Details are given in Section 3.5.) This signal is generally added to the "main" envelope signal.

These units have proved very useful for providing a variety of attack and decay characteristics. The outputs of these units can be added or subtracted in order to construct more sophisticated parameter contours. On the other hand, prior to experimenting with this equipment, I had hoped that, as a by-product, it would be possible to simulate the attack characteristics of some of the ordinary musical instrument tones. Thus far, it does not seem to be possible to do so with any degree of accuracy. Hence, it must be concluded that other sub-parameters which have not been taken into account by this system are needed to describe these tones.

In several of the mathematical expressions which follow the convenient mathematical "step function" U(x) is used. It is given by

$$U(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases} \qquad (3\text{-}1)$$

3.2 Time-Envelope and Gate Control

Simple modular control circuitry is now in operation which converts a voltage step input of either polarity and of at least eight volts magnitude (i.e., $V_{in} = V\,U(t)$, where $V \geq 8$) to a set of six outputs each having the form

$$V_{out} = \pm 9A\,U(t')\,[1 - e^{-t'/\tau_2}], \qquad \text{(Attack Mode)}; \qquad (3\text{-}2a)$$

or

$$V_{out} = \pm 9A \cdot U(t') \left\{ e^{-[1/2\tau_2 - \sqrt{1/(2\tau_2)^2 + 1/\tau_3^2}]t'} - e^{-[1/2\tau_2 + \sqrt{1/(2\tau_2)^2 + 1/\tau_3^2}]t'} \right\} e^{-t'/\tau_3}$$

$$\doteq \pm 9A\,U(t')(1 - e^{-t'/\tau_2})\,e^{-t'/\tau_3} \quad \text{for } \tau_2 \ll \tau_3, \text{ (Decay Mode)}; \qquad (3\text{-}2b)$$

where $t' = t - \tau_1$ and t is the time variable.

τ_1, the delay time, can be varied continuously in three ranges: 0 to 50 ms., 0 to 500 ms., and 0 to 5 seconds. The attack time, τ_2 varies in steps of 0, 2, 3, 5, 10, 20, 30, 50, 100, 200, 300, 500, and 1000 ms.; the decay time, τ_3, in steps of .02, .04, .08, .12, .2, .4, .8, 1.2, 2 and 4 seconds. A is the relative amplitude of the envelope signal and can be adjusted between 1 and .1. Switches are provided for setting the outputs to either polarity and to either the Attack or the Decay Mode.

An input amplifier circuit accepts two DC inputs, one positive and one negative. If either of the two inputs exceeds 8 volts in absolute magnitude, the output of the amplifier is +12 volts, whereas if they are both zero, the output is zero. The output of the amplifier is then applied in parallel to the inputs of six "Time-Delay Gate (T.D.G.) Networks". The function of each of these networks is to produce a step voltage output which is time-delayed by an adjustable amount with respect to the input and to turn on an arbitrary external signal in synchronism with the step voltage output. When the input to these networks is zero, the outputs are also zero and the external signals are turned off. The schematic for one T.D.G. Network is given by Figure 3.2.

The R-C network at the input of the first transistor stage of the T.D.G. Network is arranged so that the input current builds up at a rate determined by a variable resistor and a shunt capacitor until this current reaches a threshold value and the two relays simultaneously are turned on. One relay is used to switch the output from zero volts (ground) to either positive or negative 18 volts, and the other is used to turn on the external signal. When the input voltage drops to zero, current conducted by the silicon diode in shunt with the variable resistor rapidly discharges the capacitor, and the relays return to their normal positions. The time delay is actually quite independent of the input voltage, providing it is equal to or greater than 12 volts, because the zener diode limits the R-C network input voltage to this value. Resistors R_1, R_2, R_3,

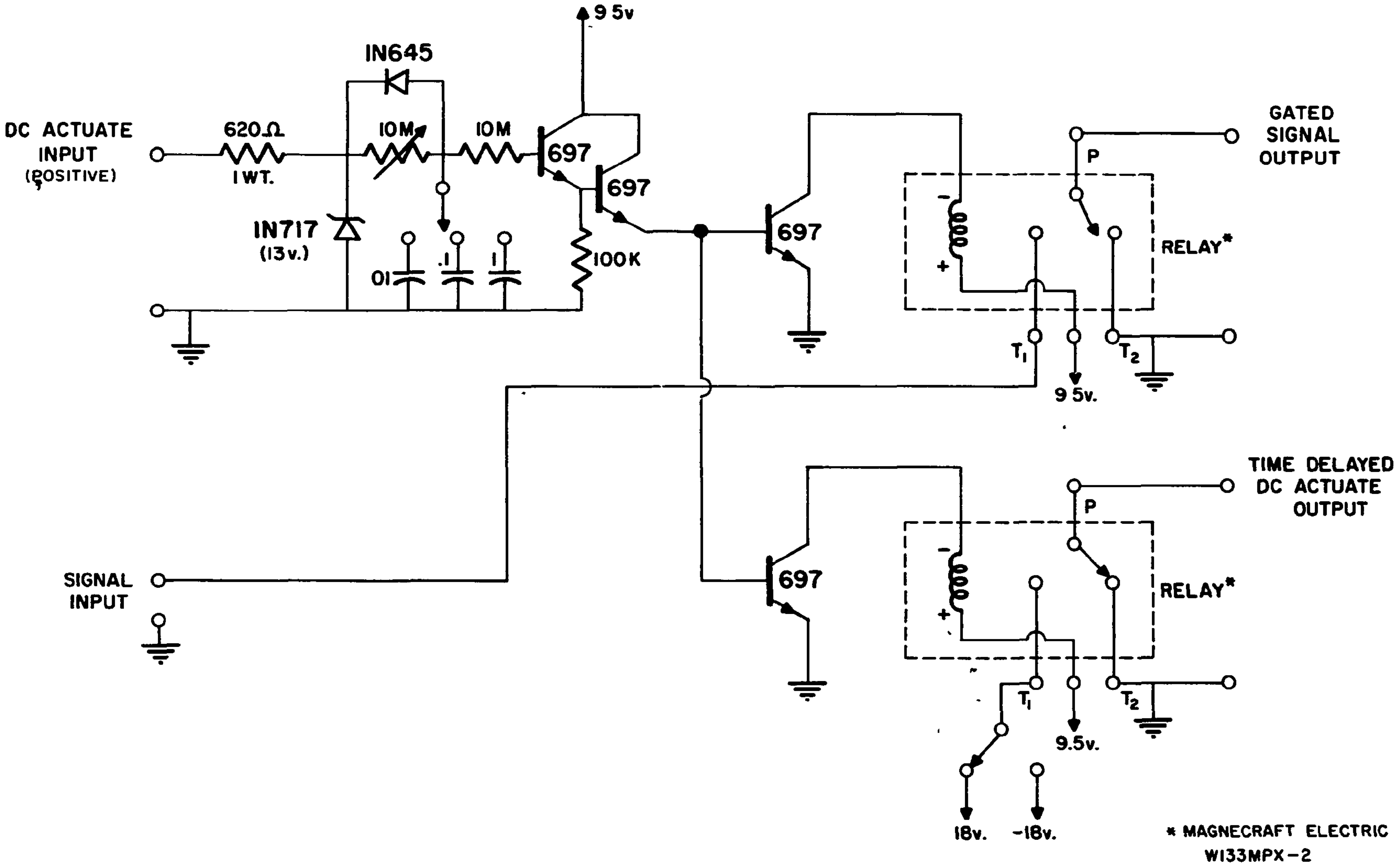


Figure 3.2 Time-Delay Gate Network Schematic.

and R_4 are adjusted to equalize the firing times of the two relays and to set the maximum time delay to exactly 5 seconds. The time delay due to the action of the relays themselves is approximately 2 ms.

The step voltage output of each T.D.G. Network is normally applied to the input of one of six "Time-Envelope (T.E.) Networks", the schematic of which is given by Figure 3.3. Note that the networks can be set to Attack or Decay modes for either positive or negative step voltage inputs. Again diodes are used to speed up the discharge of the capacitors so that the networks will respond correctly to repetitive actuating inputs. The proper load resistance for these networks is 100 kohms, which is the standard input impedance of the Harmonic Tone Generator. Assuming that it is driven by a T.D.G. Network, the general forms for the output of a T.E. Network for both the attack and decay modes are given by Equation (3-2). Normally, if A = 1, the attack time is $\tau_2 = 2x10^5 C_2$. The complexity of the expression for the Decay Mode output envelope signal, as given by Equation (3-2b), is necessary for an accurate definition in terms of these values for τ_2 and τ_3. The variable series resistor connected to the output serves to attenuate the output signal by a factor not exceeding 10. Simultaneously, the attack time is increased to a maximum of 1.8 times its normal (zero attenuation) value, and the decay time is increased to a maximum of 5 times its normal value.

Thus far, I have described only one out of the several input-output configurations of the T.D.G. and T.E. Networks which are possible. When all the inputs of each of the individual T.D.G. Networks are connected to the output of the input amplifier, the T.D.G. Networks are in "Parallel Mode." However, a switch is provided to place these networks in "Series Mode", whereby network

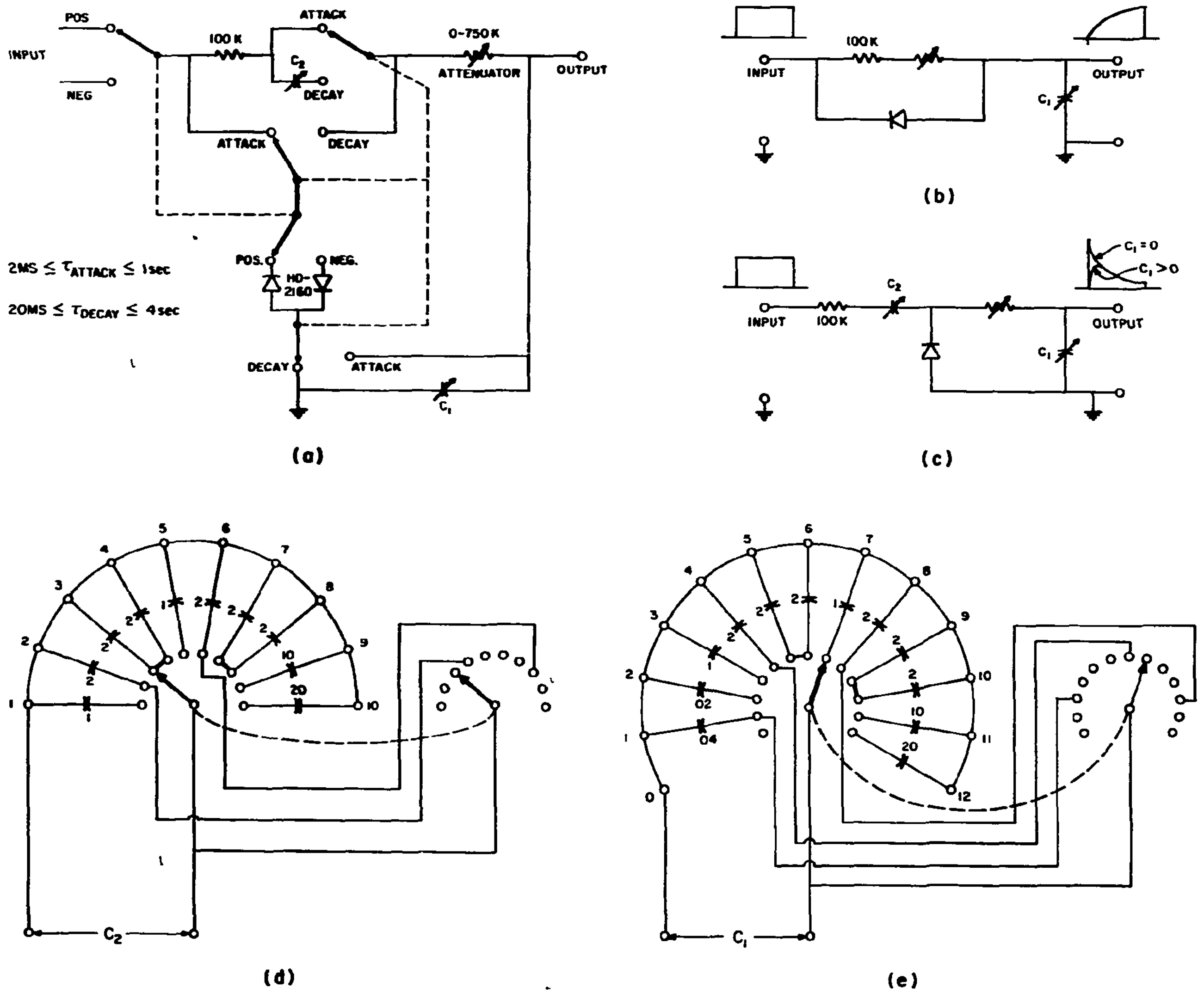


Figure 3.3 Time-Envelope Network Circuit Diagrams: (a) Basic Network (b) Attack Configuration (c) Decay Configuration (d) Circuit for Switching C_2 (Decay) (e) Circuit for Switching C_1 (Attack)

1 is driven by the input amplifier, but network 2 is driven by the voltage step output of network 1, network 3 by network 2, and so on. Either one or the other of these two modes may be the best one for accurately producing a rhythmic sequence of time delays, depending on the particular spacing of the sequence. The T.D.G. Networks can also be set in "Independent Mode", in which case the individual networks can be actuated directly by six independent sources. Finally, they can be set in "AC Mode", whereby the outputs of two AC-DC converters are connected to the inputs of two independent sets of three T.D.G. Networks (the AC-DC converter is described in Section 3.4).

In a similar fashion, the Time-Envelope Networks can be set in "Time-Delay Output Mode", as described directly above; or "DC Actuate Mode", in which case the T.E. Networks are connected directly in parallel to the input amplifier inputs, i.e., the T.D.G. Networks are by-passed. The T.E. Networks may also be set in "AC Mode" or "Independent Mode" in which cases the inputs to these networks are exactly the same as those given for the corresponding modes of the T.D.G. Networks.

The two inputs of the input amplifier can be set at any one of three positions: (a) "External Source", (b) "Manual" (a pushbutton is provided to actuate the system), or (c) "Keyboard Output" (described in Section 3.3).

The six independent envelope signals may be applied to the Harmonic Tone Generator amplitude control inputs in order to program a completely different attack-decay characteristic for each harmonic. Also, since each harmonic amplitude is proportional to the sum of three separate control inputs to the Harmonic Tone Generator (See Figures 2.2 and 2.9), it is possible to program amplitude contours which are quite different from that depicted by Equation (3-2). For example, the following envelope characteristic can be programmed using two Attack Mode envelope signals of opposite polarity:

$$U(t)\left[1 - e^{-t/\tau_{2a}}\right] e^{-(t - \tau_{1b})U(t - \tau_{1b})/\tau_{2b}}$$

$$\doteq U(t)\left[1 - e^{-t/\tau_{2a}}\right] - U(t - \tau_{1b})\left[1 - e^{-(t - \tau_{1b})/\tau_{2b}}\right]$$

$$\text{for } \tau_{1b} \gg \tau_{2a} . \qquad (3\text{-}3)$$

The expression on the left side of Equation (3-3) represents a signal which is characterized by an attack at time zero of duration τ_{2a}, a steady state which lasts until t equals τ_{1b}, and an exponential decay which begins at $t = \tau_{1b}$ with decay time constant τ_{2b}. As is indicated by the right side of Equation (3-3), this more complex envelope contour can be derived from the addition of two simple attack envelopes, the second being of opposite polarity and time-delayed with respect to the first. Other interesting envelope characteristics can be similarly constructed. An "overshoot" characteristic can be obtained by the addition of a short decay envelope to one of longer duration.

When applied to the frequency control input, one of these signals will produce a frequency glide or portamento. In conjunction with an analog multiplier (as described in Section 3.5) they can also be used to "shape" musical parameters such as vibrato and tremolo.

3.3 Keyboard Pitch and Rhythm Control

A convenient source of a pitch control voltage and simultaneously produced actuate voltage is an electronic organ keyboard which has been wired for a series of adjustable pitches. When a particular key is depressed, a switch is closed making contact between a fixed reference voltage and one of a large parallel group of variable resistors (one for each key) which are all terminated by a fixed resistance (see Figure 3.4). The variable resistor of each voltage divider thus created is adjusted so that the output voltage, which is applied to the Harmonic Tone Generator frequency control input, determines that the frequency correctly corresponding to this key is generated by the Harmonic Tone Generator. The proper output voltage is determined by Figure 2.5. The pitch control voltage varies from -.7 to -11.5 volts corresponding to pitches between C_2 and E_7 when keys are depressed, but is zero when no key is depressed. When playing this type of "parallel tuned" keyboard, the performer must be aware that two keys depressed will produce a pitch higher than both of the pitches corresponding to each of the keys.

A +12 volt actuate voltage is derived from the pitch control voltage by means of the limiter-amplifier circuit as shown. The operational amplifier is set for a gain of $12/-.7 = -17$ causing the output voltage to limit at about 12 volts, its maximum output. Nevertheless, zero pitch control voltage produces a zero actuate voltage within a close tolerance. In this way both the pitch control voltage and the actuate voltage are produced from single contacts and are automatically synchronized. The actuate voltage, normally used to initiate attacks, must be reset to zero by lifting each key before depressing the next one. The capacitors in the T.D.G. and T.E. Networks must be discharged before a subsequent transient can be produced. Strict legato with no attack transients and constant intensity control can be performed only with the T.E. Networks set on Attack Mode.

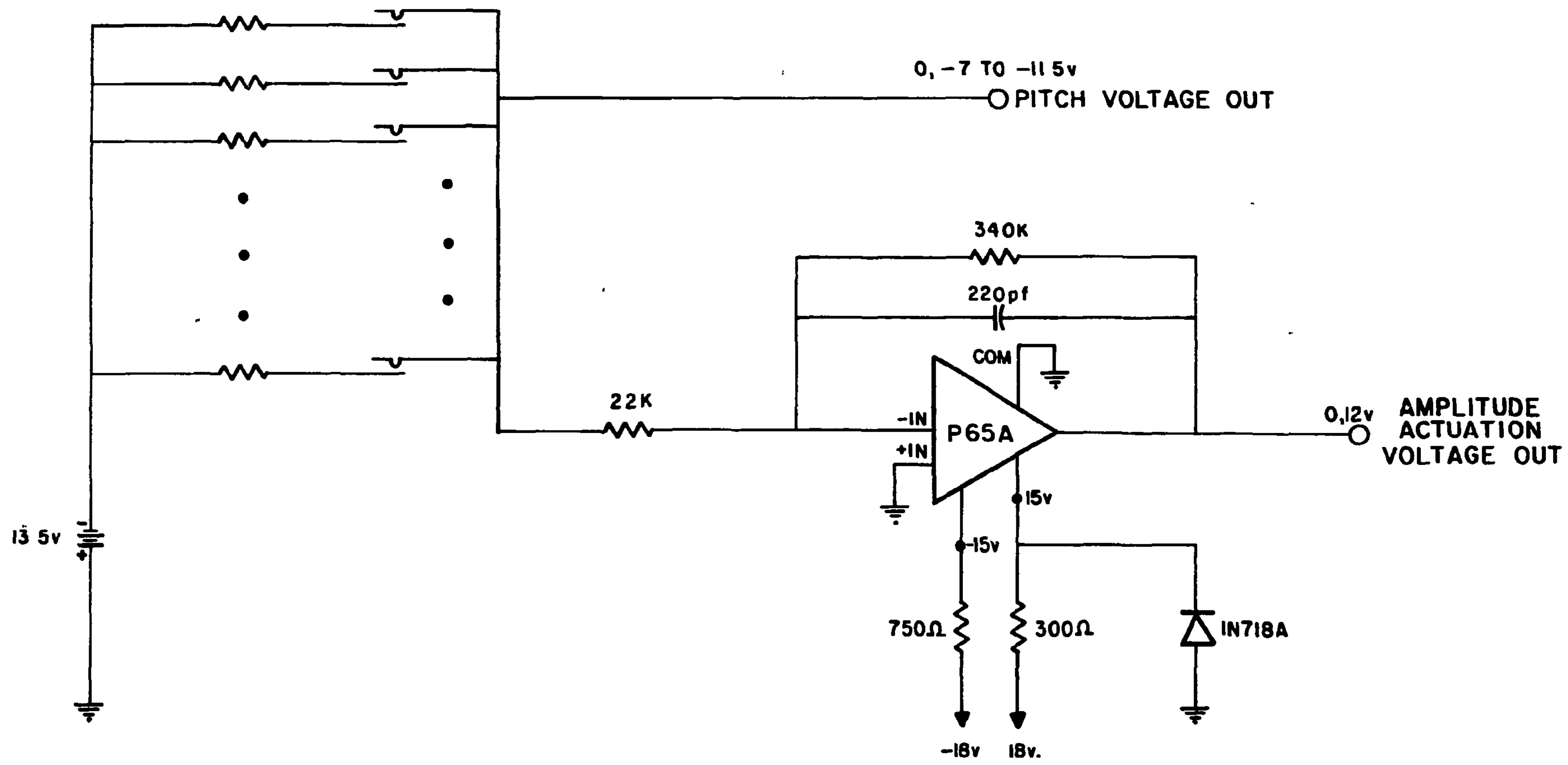


Figure 3.4 Keyboard Schematic.

3.4 Rhythm Control by Tape and AC-DC Converter

A circuit which converts tone bursts whose frequencies are 5 kc or greater into DC actuate voltages has also been constructed, and another identical circuit is to be completed shortly. Two simultaneous sequences of tone bursts recorded on two-channel tape will be then used to actuate two sets of time delay and envelope controls, or they can be used to gate sounds according to the rhythmic pattern of the tone bursts. The circuit, as shown in Figure 3.5, consists of a limiter, gate, full-wave rectifier, low pass filter, and output amplifier. The rise time of the output voltage step is about 1 ms. By choosing the tone burst frequency to be 7000 - f_a, where f_a is the desired audio fundamental, the tone bursts can be used to simulate the Harmonic Tone Generator voltage-controlled oscillator. (An input is provided for this purpose, see Figures 2.2 and 2.9). Hence, the same pitch and rhythm sequence can be repeated for different manual settings or variations of the attacks, harmonic amplitudes, frequency modulation, and other parameters.

Rhythmic patterns of tone bursts may be recorded on magnetic tape by two methods. First, sections of tape containing a pre-recorded high frequency tone and sections of blank tape may be spliced together according to the desired pattern. In certain circumstances this method may offer a modest advantage over the usual procedure of splicing and measuring tones since it is not necessary to search out and splice at the beginnings of tones already recorded. Second, a digital computer can be programmed to produce square wave tone bursts according to an arbitrary time pattern specified as data for the program, and these tones can be recorded on magnetic tape. Composers have used tones generated by the University of Illinois CSX1 Computer at the Coordinated Science Laboratory for this purpose.

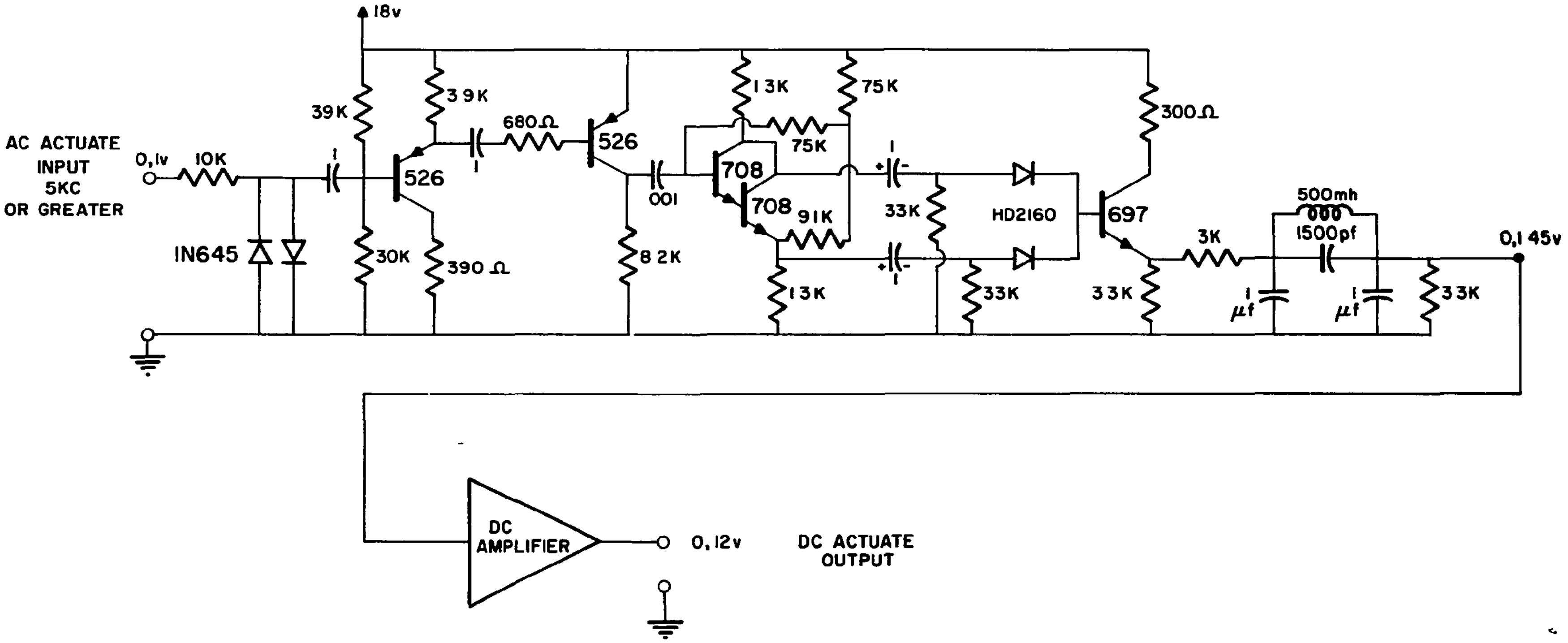


Figure 3.5 AC-DC Converter Schematic.

3.5 Analog Multiplier

A compact device which produces an output signal proportional to the instantaneous arithmetic product of two DC signals, irrespective of their polarities, was needed for the synthesis of more complex parameter contours. It should be noted that the two types of amplitude modulation devices used in most electronic music studios will not satisfactorily perform this function. One of these, the "ring modulator", is restricted to multiplying AC signals over a comparatively small dynamic range. The other type, the "voltage-controlled-gain amplifier", is restricted to the multiplication of a bi-polar signal and a mono-polar signal and, in addition, may have a non-zero DC offset voltage. Therefore, a special analog multiplier was designed and constructed which takes the two input voltages e_1 and e_2, and multiplies them to produce the output

$$e_{out} = K\, e_1 e_2. \tag{3-4}$$

This multiplier uses a carrier frequency method to derive the product of two signals. The first signal controls the amplitude of an ultrasonic frequency square wave, and the second signal controls the phase of another square wave of the same frequency. The two square waves are then compared by means of a phase detector. To be specific, one square wave with amplitude proportional to e_1 is switched on and off in synchronism with the positive and negative portions of the other square wave which drives the phase detector. This latter square wave is generally out of phase with the former, and the deviation from 90^o of the phase difference between the two is varied by the control voltage e_2. Figure 3.6a shows the output waveform (before filtering) for six combinations of values e_1 and e_2. Given that the peak-to-peak voltage of this wave form is

e_1 and the phase shift is $\pi/2 - ke_2$, it follows that the average DC component of the waveform is

$$e_{out} = ke_1e_2/2\pi . \tag{3-5}$$

The block diagram for a "phase shift" multiplier is given by Figure 3.6b.

a. A Practical Phase Shift Multiplier

Some basic requirements with respect to use of the multiplier for parameter control as well as for general electronic music purposes are the following:

1. Ideal input-output relationship· $e_{out} = K e_1e_2$.
2. Bandwidth e_1 (amplitude control): 3 kc.
3. Bandwidth e_2 (phase control): 10 kc.
4. Maximum DC input or output. $\pm$ 10 volts.
5. AC feed-through with $e_1 = 0$. 70 db below maximum output.
6. AC feed-through with $e_2 = 0$ 40 db below maximum output.
7. Maximum DC offset voltage at output: 10 mv.
8. Distortion of input sine wave when other voltage is set to 10 volts 1 per cent maximum.
9. Input impedances· 100 kohms.
10. Outputs: one DC output drives 10 volts into 10 kohms and one AC output drives 5 volts rms. into 600 ohms.

The heart of the phase shift multiplier is the pulse shaping circuit shown in Figure 3.7. The output of the integrated chopper at the top of the diagram is a square wave alternating between 0 and $-e_1$, and the voltage developed between the outputs of the top and bottom choppers is a balanced square wave alternating between $-e_1/2$ and $+e_1/2$. At the input to the output differential

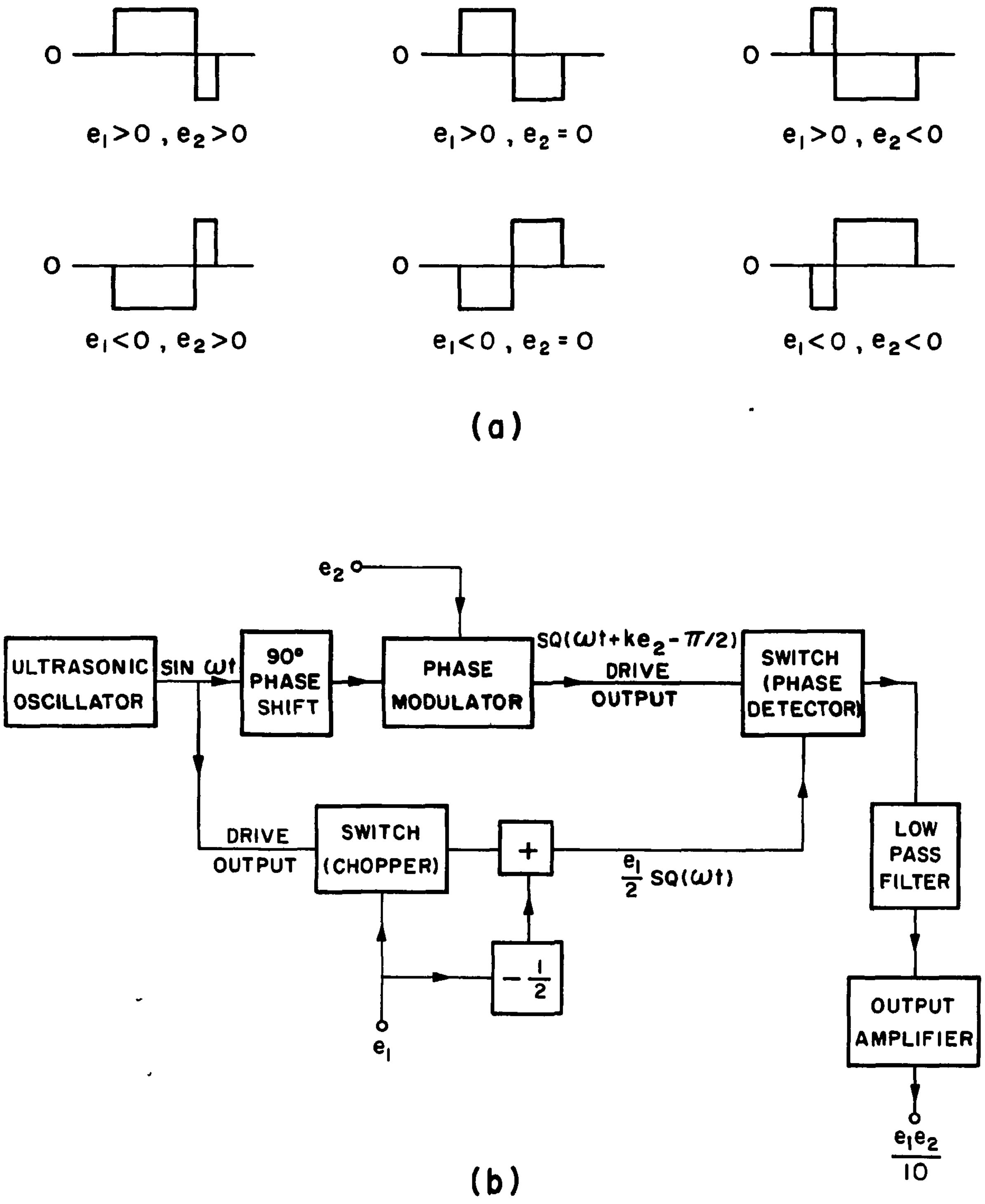


Figure 3.6 Phase Shift Multiplier: (a) Output Waveforms (Before Filtering) (b) Block Diagram.

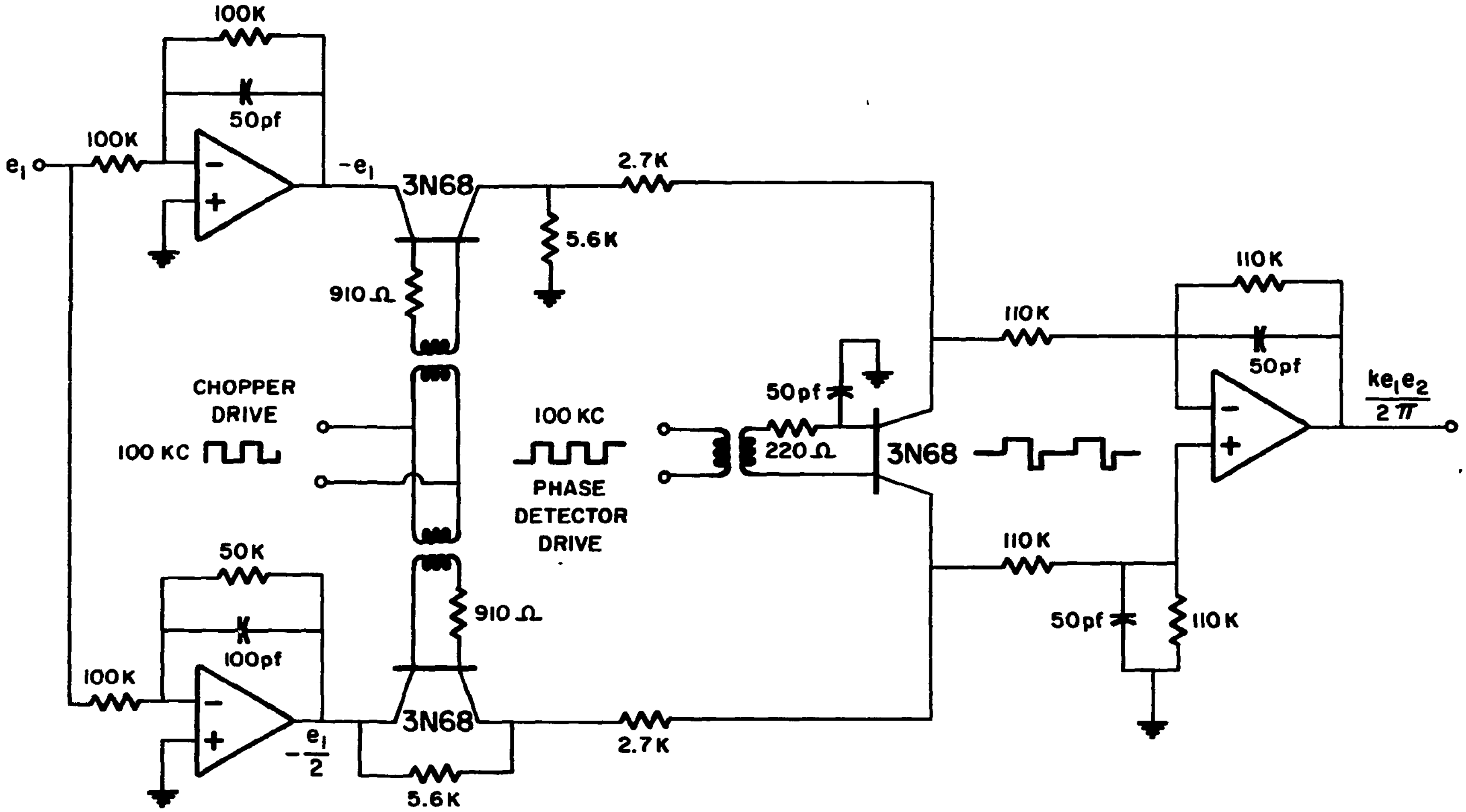


Figure 3.7 Phase Shift Multiplier: Pulse Shaping Circuit Diagram.

amplifier the voltage across the phase detector chopper alternates between this square wave and 0 volts according as the phase detector is positive or negative, respectively. Hence, the waveforms depicted in Figure 3.6a result and the low frequency output of the amplifier is $ke_1e_2/2\pi$. The bottom chopper is not used to switch the signal but is used for error current cancellation. When e_1 is zero, a series of "spikes" remains at the output of both choppers. However, since in this mode the chopper circuit configurations are identical, the spikes are essentially cancelled across the phase detector. This results in an improvement of the AC feedthrough with $e_1 = 0$ by about 20 db.

The method used for phase modulation is illustrated by the block diagram shown in Figure 3.8. The output of a 100 kc carrier frequency oscillator is AM-modulated by means of a chopper circuit (see Figure 2.8) and a second order Butterworth band pass filter with 20 kc bandwidth to produce an in-phase sine wave with amplitude $e_2/2$. The oscillator output is also shifted by 90° to produce a cosine wave with a fixed 2.5 volt amplitude. When these two quadrature components are added, the phase of the resulting sinusoid is equal to $\pi/2 - \tan^{-1}(e_2/5)$ relative to the phase of the oscillator output. However, if the magnitude of e_2 is less than 2 volts, this relationship is linear within 2 per cent, i.e., the phase is approximately equal to $\pi/2 - e_2/5$. [The linear range for e_2 could be extended to ± 10 volts if the buffer amplifier were preceded by a diode function generator designed to deliver the output $5\tan(e_2/5)$]. This relationship is also inherently symmetrical about $\pi/2$ for positive and negative values of e_2. It follows that k in Equation (3-5) is 1/5, and the output of the differential amplifier shown in Figure 3.7 is

$$e_1\tan^{-1}(e_2/5)/2\pi \doteq e_1e_2/10\pi, \text{ for } e_2 < 2. \tag{3-6}$$

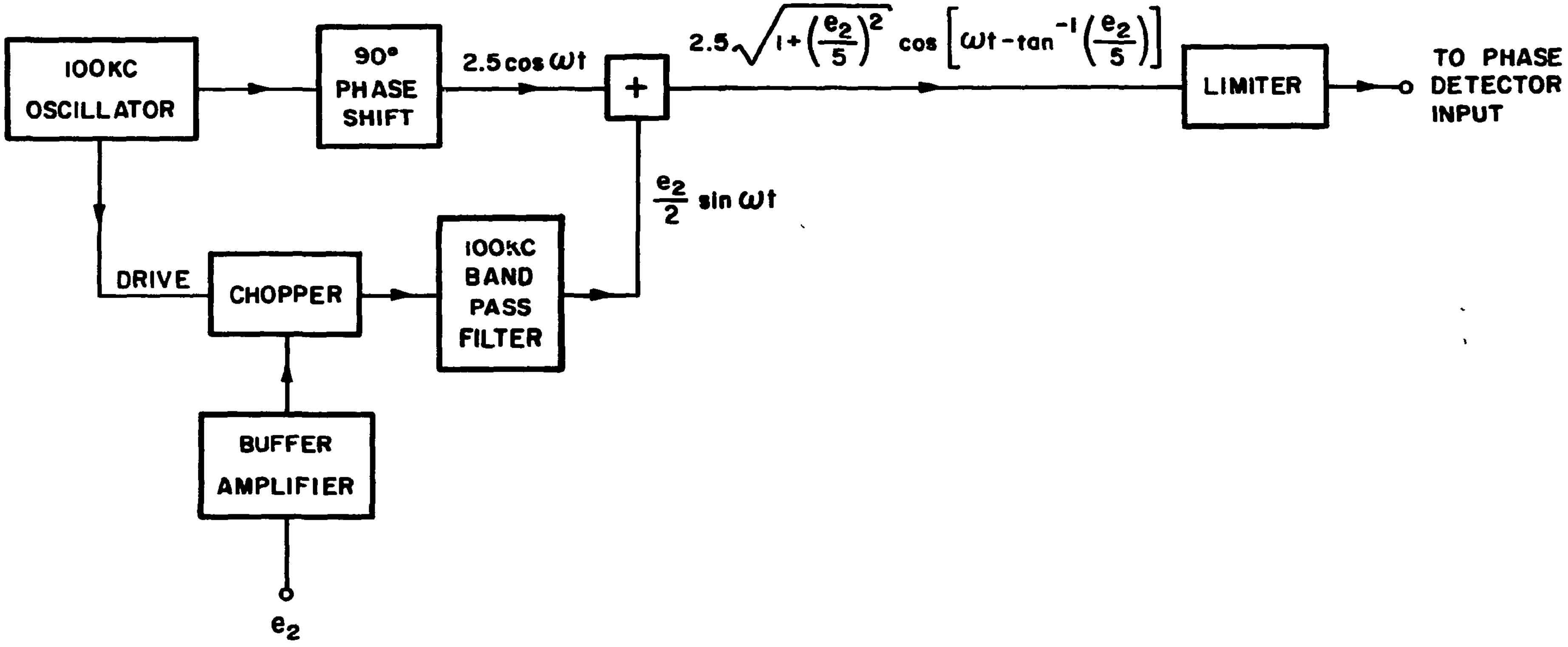

100KC OSCILLATOR
90° PHASE SHIFT
2.5 cos ωt
+
$2.5\sqrt{1+\left(\frac{e_2}{5}\right)^2}\cos\left[\omega t-\tan^{-1}\left(\frac{e_2}{5}\right)\right]$
LIMITER
TO PHASE DETECTOR INPUT
$\frac{e_2}{2}\sin\omega t$
DRIVE
CHOPPER
100KC BAND PASS FILTER
BUFFER AMPLIFIER
e_2

Figure 3.8 Phase Shift Multiplier: Phase Modulator Block Diagram.

However, it was decided that the magnitude of the output of the entire multiplier should be 10 volts when both inputs are ± 10 volts. Hence, the gain of the output amplifier (Figure 3.6) is adjusted to $2\pi/\tan^{-1}(2)$, so that the relationship given by Equation (3-4) is valid for $K = 1/(5\tan^{-1}(2))$ and for $|e_2|$ less than 2 volts.

Unfortunately, two anomalies of the circuitry and the technique limit the performance of the constructed multiplier with respect to satisfying requirement 6, namely, the AC feed-through with $e_2 = 0$, especially when e_1 contains higher frequencies. The performance was improved by compensating for these anomalies at the input of the multiplier output amplifier. The maintenance of zero DC output when $e_2 = 0$ depends on the cancellation of the areas of the positive portion and the negative portion of the pulse wave [See Figure 3.6a center]. The response of a simple integrated chopper circuit [Figure 2.8] to both +10 and -10 volt DC control voltages was measured, the rise and fall times of the output square wave for the positive control voltage were 320 ns. and 400 ns., respectively, whereas for the negative control voltage they were 200 ns. and 150 ns., respectively. The result of this asymmetrical response was that the output for $e_2 = 0$ had roughly the form

$$\epsilon_1 e_1 + \epsilon_2 |e_1| \tag{3-7}$$

This output error was partially cancelled by the introduction of opposite polarity voltages corresponding to these two terms.

Also, when e_1 contains higher frequencies, the approximation that the pulse waves derived from sampling e_1 are square is no longer valid. The error in cancellation increases with increasing frequency of e_1. Therefore, an opposite polarity version of e_1 which increases in amplitude with frequency (obtained by feeding the signal through a capacitor of optimum value) was used to compensate this error at the output, and a substantial improvement resulted.

Some performance characteristics of the practical multiplier which have been measured are:

(1) AC Feed-through

frequency (cps.)	A(db)	B(db)
50	-75.5	-42
500		
1,000		
2,000	-58	
3,000	-55	
4,000	-53	
5,000	-52	
10,000	-51.5	

Conditions:

A. $e_1 = 0$, $e_2 = 2$v.rms.
reference to $e_1 = \pm 10$ v., $e_2 =$ 2 v.rms.

B. $e_1 =$ 2 v. rms., $e_2 = 0$
reference to $e_1 =$ 2v.rms., $e_2 = \pm$ 10v.

(2) Harmonic Distortion (f=50 cps., $e_1 = +10$ v.)

e_2(volts rms)	% distortion at output
.5	.54
1.0	.68
1.5	1.1
2.0	1.6
2.5	2.5
3.0	3.4

b. Parameter Control by Multiplication

An envelope signal may be considered to consist of two terms: (1) a smooth "average" envelope function $A_1(t)$ such as that given by Equation (3-2), and (2) a tremolo perturbation function $p(t)$ with zero average value. An envelope characterized by a changing percentage perturbation can be produced by multiplying a suitable envelope function $A_2(t)$ with a readily available steady-state low frequency signal $s(t)$ having a constant rms. amplitude (such as a sine wave or filtered white noise). Hence, the total envelope signal in this case is

$$A_T(t) = A_1(t) + p(t) = A_1(t) + A_2(t).s(t). \tag{3-8}$$

For example, the envelope for a simulated "vibraphone" tone can be constructed by taking $A_1(t) = 0$, $A_2(t) = U(t)e^{-t/\tau}$, and $s(t) = \sin(2\pi ft)$ where τ and f are the order of 1 second and 5 cps., respectively·

$$A_T(t) = U(t)e^{-t}\sin(10\pi t). \tag{3-9}$$

Analogously, the multiplier may be used to vary the amount of vibrato with respect to time. In this case, the "main" pitch control voltage is complemented with a vibrato perturbation signal of the same type depicted above. As an example, suppose that the Harmonic Tone Generator pitch control voltage is given by

$$V_f(t) = V_o + A_2(t) \cdot s(t). \tag{3-10}$$

Then, providing $A_2(t)\cdot s(t) = p(t)$ is sufficiently small, the instantaneous fundamental frequency at the output of the Harmonic Tone Generator is given by Equation (2-1) to be

$$f(t) = q[V_f(t)] \doteq q(V_o) + \frac{\partial q}{\partial V_f} \cdot A_2(t)\cdot s(t). \tag{3-11}$$

If $A_2(t)$ is an increasing function with respect to time, a tone can be created which is initially "flat" and is gradually transformed into a tone with definite vibrato.

c. Other Uses for Multipliers in Electronic Music

Since the inception of electronic music, the transformation of spectrums from one pitch region to another has played an important role in electronic music composition. Usually this has been accomplished by the so-called "ring modulator", a multiplier of AC signals which derives its name from its circuit configuration--a ring of four diodes.[24] A sound, having a simple, changing harmonic spectrum and a distinctive attack-decay characteristic, which is multiplied by a pure sinusoid can be transformed into a sound with a more complex, and usually inharmonic, spectrum retaining the original attack-decay pattern. If the spectrum of the original sound contains the frequencies f_1, $2f_1$, $3f_1$, ... and the frequency of the sinusoid is f_2, the output of the multiplier is a double side-band spectrum consisting of the frequencies ..., f_2-f_1, f_2-2f_1, f_2+f_1, f_2+2f_1,... The amplitudes of these frequencies are proportional to those of the corresponding harmonics of the original signal. Under these circumstances, it is essential that the frequency f_2 not be present in the output spectrum since its constant value represents an undesirable aural redundancy. The appearance of the "carrier" frequency f_2 may be due to an inbalance in the modulation with respect to the original sound or by a direct feed-through of f_2 to the output. Therefore, an analog multiplier can accomplish spectrum transformation satisfactorily only if it is properly balanced with respect to these two items and has adequate frequency response.

---

[24] Wilcox, R.H. · "Use of a Diode Ring as a Four Quadrant Multiplier," Review Scientific Instruments, 30, 1009-1011, 1959.

The multiplier may also be used for the instantaneous amplitude control of an arbitrary sound signal, where the amplitude control voltage is supplied by one of the Time-Envelope networks described in Section 3.2. Here it is important that the modulation be linear with respect to the original signal and that none of this signal feed through to the output uncontrolled. As discussed in Section 2.4c, a large dynamic control range, at least 70 db, is needed to reduce the signal feedthrough (or "leak") to a tolerable level.

d. Some Other Possible Designs for a Electronic Multiplier

It is ironic that a signal accurately proportional to the sum of two electrical signals can be produced by so simple a circuit as the mere connection of two resistors, and yet the generation of an accurate product of the two signals inevitably requires a much larger number of components. Analog computation usually depends on the availability of natural devices which follow simple mathematical laws (a case in point: the capacitor as an integrator of current), but there does not seem to be in nature a simple, inexpensive, accurate device for multiplication. At this point, it seems desirable to review some other possible methods of electronic multiplication, besides the "phase shift" method, which were considered.

Probably the most popular electronic multiplier design uses the "quarter-square" method. This type of multiplier consists of two squaring circuits and sum and difference amplifiers; hence, the output is given in terms of the input by the expression

$$e_1 \cdot e_2 = 1/4 \, [(e_1+e_2)^2 - (e_1 - e_2)^2] \qquad (3\text{-}12)$$

A square law characteristic may be constructed as a linear-segment approximation to a parabola according to a diode function generator design[25] or by taking advantage of the natural square law behavior of certain electronic devices.[26,27] The latter technique is hampered by the fact that none of the several physical devices used have true square law characteristics over a sufficiently large dynamic range. With the former technique, large dynamic range and accurate performance can only be achieved if large numbers of diodes and precisely adjusted resistors are used. Also, in this case, the derivative of the approximation to a parabola contains many discontinuities which for certain types of input signals can result in the superposition of an undesirable high frequency "noise" on the output signal. Both techniques involve stringent balancing requirements.

Control voltages can be used to vary the gain of amplifiers or to vary voltage-sensitive resistive components (such as the field-effect transistor biased in the "pinch-off" region). Here again, dynamic range and balancing requirements restrict this technique for four-quadrant operation.

In recent years, progress has been made in ultilizing magnetic fields to obtain multiplier action in conjunction with the Hall-effect and the magneto-resistance effect in semiconductors.[27] For example, the output voltage of the Hall-effect multiplier is proportional to a flux density B and a Hall element current I_1, while B is proportional to the current I_2 in a coil which produces the magnetic flux. However, the main difficulty with this method is that in order to maintain a constant flux density over a wide frequency range, either the coil drive voltage must be very large for high frequencies (when large inductances are used) or the constant coil drive current must be very large (when small inductances are used). This means that the maximum required signal power at the coil input is proportional to the highest signal frequency. At present, these power requirements are too severe for wide band applications where a reasonably large output signal is required (1 volt). Hence, the magnetic multipliers are probably best suited for low frequency or power applications.

---

[25] Marshall, B.O. : "An Electronic Multiplier", Nature, 167, 29, 1951.

[26] Berger, F.B., MacRae, D.: "Mathematical Operations on Waveforms-II", Waveforms, chap. 19, 679-686, M.I.T. Radiation Laboratory Series, Vol. 19.

[27] Hilsum, C.: "Multiplication by Semiconductors," Electronic Engineering 30 (369), 664-666, 1958.

Carrier frequency methods of multiplication seem more promising, since techniques for nonlinear approximation and simulation are replaced by techniques for the control of sawtooth waveform linearity and the areas of square pulses.[28] The average DC component contributed by a square pulse wave is equal the product of the pulse height and the pulse width divided by the pulse period (a constant). Hence, if voltages e_1 and e_2 control the height and width of a recurrent square pulse, respectively, the average component is proportional to their product. (The high frequency pulse components are removed by a low pass filter). This can be accomplished by means of an electronic switch which turns on the voltage e_1 only when the switch actuating voltage is positive. If the sum of a sawtooth or triangle waveform (normally entirely negative) and the voltage e_2 is applied to the actuate input, the switch is turned on for a time proportional to e_2, for positive values of e_2. This is inherently a two-quadrant multiplier, although a four-quadrant multiplier can be formed either by positive biasing of e_2 at the switch actuate input, forming the product $e_1(k+e_2) = ke_1 + e_1e_2$, and subtracting ke_1 at the output, or by applying $-e_1$ and $-e_2$ to an otherwise identical electronic switch and adding the outputs of the two switches.

---

[28] Kraicer, A.: "Two Four-Quadrant Electronic Multipliers," Radio and Electronic Engineer, 25(1), 5-11, 1963.

A similar type of carrier frequency multiplier, which is inherently a four-quadrant device, is a "phase shift" multiplier. A phase shift multiplier was built for use in the parameter control system and is described in Section 3.5a.

Chapter 4

VARIABLE FORMANT FILTER (PROGRAM EQUALIZER), THEORY AND DESIGN

4.1 Introduction

A variable formant filter was designed and constructed, which is capable of assuming a large variety of shapes (frequency transfer characteristics) over a fixed frequency range. It can be used for studies of the audible effect of harmonic tones with different formant characteristics. The frequency range of the filter is partitioned into a number of band regions which are logarithmically spaced on the frequency scale. Ideally, formant curves are to be approximated in staircase fashion by altering the intensity of each band to coincide with the corresponding point on the formant curve. Physically, this is accomplished by a device containing a set of band pass filters whose center frequencies are logarithmically spaced (a fixed number per octave of frequency) and whose transfer characteristics are logarithmically symmetrical and identical when scaled to the same center frequency. The signal to be filtered is applied in parallel to each of the filter inputs, each filter output is connected to the input of a variable attenuator, and the total output of the device is formed by the sum of the individual attenuator outputs. (See Figure 4.1.) Electronic instruments of this type have been used in the broadcasting industry for some years and are called "program equalizers". However, data for the design of such instruments do not seem to be available.

As I discussed in Section 1.5, a voltage-controlled subtractive synthesizer can be formed by the connection of the output of a voltage-controlled oscillator, whose waveform contains a large number of harmonics, to the input of a formant

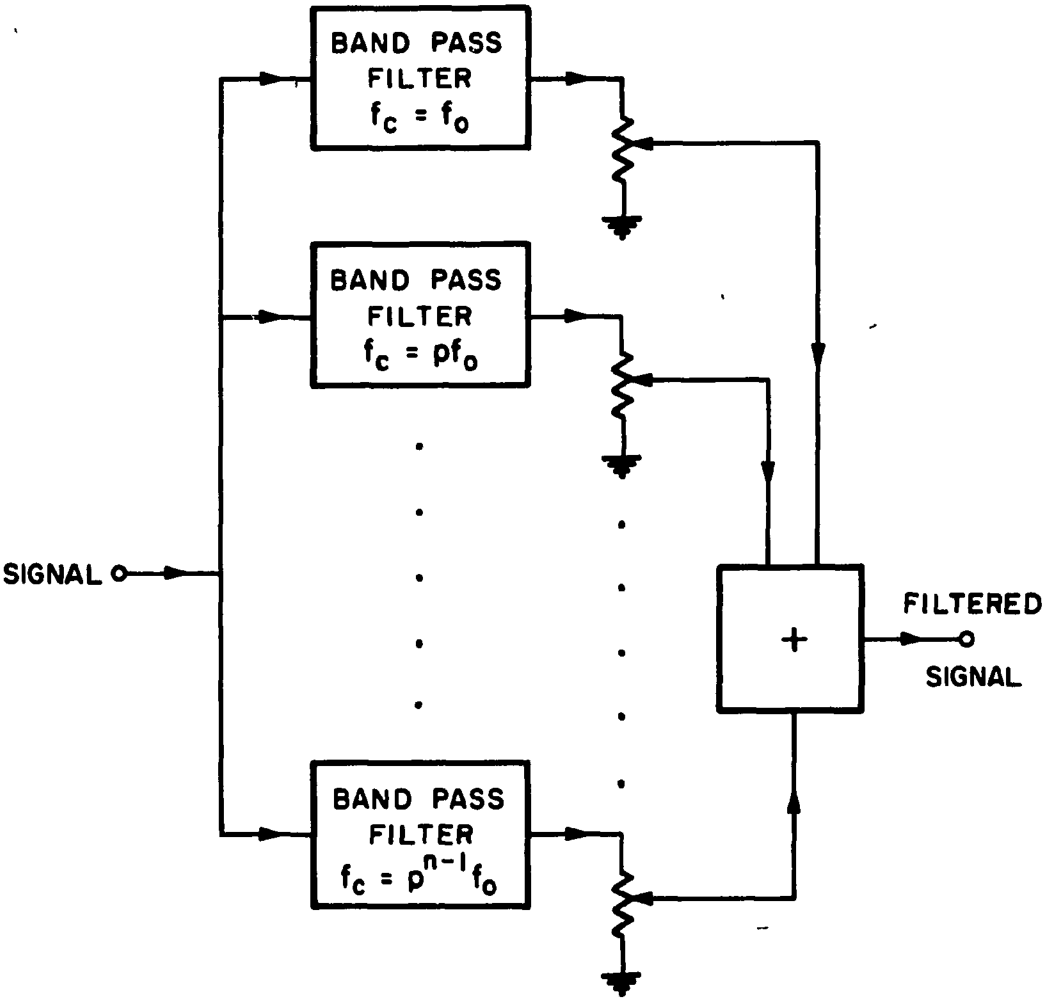


Figure 4.1 Variable Formant Filter Block Diagram.

filter, whose shape can be varied by means of control voltages. The shape control can be implemented using the variable formant filter just described, since this device can be easily modified for the functional replacement of the variable attenuators with voltage-controlled amplifiers. Fortunately, Robert Moog has recently designed and made commercially available voltage-controlled oscillators and amplifiers for use in the electronic music studio.[29] Moog's oscillators provide a choice of sine wave, square wave, pulse wave, sawtooth wave, and triangular wave outputs, which are basic waveforms for subtractive synthesis Also, the voltage-controlled amplifiers in conjunction with time-envelope voltages supplied by the equipment described in Chapter 3 can be used to provide attacks on sounds produced by subtractive synthesis. Hence, these units complement very nicely the additive synthesis system described in Chapters 2 and 3. The Experimental Music Studio has acquired some of these units and they are presently being used in conjunction with the Variable Formant Filter to be described in this chapter.

It is evident that the number of degrees of freedom for synthesizing a given frequency characteristic with a variable formant filter is proportional to the number of filters per octave (i.e., the filter density) which is employed. However, economic considerations and the need for a wide audio frequency range (40 cps. to 16,000 cps.) limit the practical density of filters to a low number per octave. In this chapter I will concentrate on finding the "best" low-order band pass response characteristic which is compatible with a given filter density.

---

[29] Moog, R.A. : "Voltage-Controlled Electronic Music Modules," J. Audio Eng. Soc., 13(3), 200-206, 1965.

The calculation of the composite response of the formant filter is complicated due to the necessity of taking into account the phases as well as the amplitudes of the band pass filter outputs.

4.2 A"Best" Criterion for Formant Curve Simulation

Arbitrarily, a criterion has been chosen such that the response of the adjustable formant filter corresponds closely as possible to that which is implied by the settings of the attenuators for two extreme cases:

(1) All attenuators set for unity gain (0 db); then the output should be flat over the range of the filter system.

(2) A single midband attenuator is set for zero gain ($-\infty$ db); in this case, the output at the corresponding midband frequency should be zero.

These conditions correspond to characteristics which are useful in musical applications: (1) Equal transmission of all frequencies ("all-pass" characteristic). (2) Rejection of one or more discrete frequencies ("notch" characteristic). The third characteristic which is useful--boosting of a single frequency--is inherent in the type of system chosen. The maximum amount of boost at one frequency relative to the others is, of course, obtained when all the attenuators are set to zero gain, except the attenuator connected to the filter passing this frequency, which is set for unity gain.

The formant characteristics of ordinary musical instruments are characterized to a high degree by resonances and antiresonances, i.e., distinct maxima and minima. Therefore, from the standpoint of their simultaion, it is advantageous to have a system capable of producing sharp peaks and notches at discrete frequencies. Also, since each value of the overall frequency characteristic varies continuously as a function of the positions of the attenuators (with those attenuators corresponding to the filter band centers nearest to this value having the largest influence), it is evident that an infinite variety of formant characteristics can be simulated and, to an accuracy of one bandwidth and with a certain restriction on the maximum slope of the curve, any desired formant shape can be simulated. A qualitative notion of the maximum slope restriction is indicated by the peak and notch characteristics of the filter set.

4.3 Computation of Ripple and Notch Characteristics of an Infinite Set of Band Pass Filters

The mid-region frequency response of a large set of band pass filters with unattenuated outputs summed together is very close to that of an infinite set of such filters, assuming that we can neglect the contribution of the filters whose center frequencies are sufficiently removed from the mid-region. Because of the logarithmic symmetry of the filters under consideration, their congruent shapes, and their even spacing on a logarithmic frequency scale, the infinite set frequency characteristic is logarithmically periodic. The period is equal to the spacing between the filter band centers. Also, the response has mirror symmetry about the band center frequencies and the geometric means of adjacent band center frequencies. Hence, the entire response function is characterized by a single half-frequency interval between one of the band centers and an adjacent geometric mean frequency.

It is shown below that among other properties of an infinite filter set, the complex response function is real at the half-interval end points and that relative maxima or minima of the magnitude function occur at these points. Evidently, when the bandwidths of the individual filters are sufficiently small (assuming their own maxima are at band center), the total response maxima lie at the band center frequencies and the minimum values lie at the geometric mean frequencies between them. It is possible for sufficiently large bandwidths of filters of order greater than 1 (the tuned circuit) for other maxima and minima to appear. Nevertheless, over a good range of percentage bandwidth values, the extent of which is tested by a different method described later in this chapter, the maximum and minimum values can be inferred to be equal or close to the response function evaluated at a band center frequency and a geometric mean frequency. Therefore, the ripple, which describes the closeness of approximation to an all-pass characteristic, can be computed to be the difference in decibels between these two values. Likewise, the "notch" attenuation can be computed as the difference between this maximum decibel value at a band center and its value when the corresponding band pass filter is left out. The advantage of this method of determining the ripple and notch characteristics, is that for each filter type, value of percentage bandwidth (or its inverse, Q), and filter density, only two real computations are required.

In the remainder of this section of this chapter, some basic definitions are made and the fundamental properties of infinite "all-pass" filter sets outlined above are proved. The response characteristic of each of the band pass filters of the set is derivable from that of an equivalent rational low pass filter response function, f(s). Parameters which are required to define the response characteristic of the infinite filter set are (1) the coefficients of the numerator and denominator polynomials of f(s); (2) the filter density N, and (3) the band pass filter Q. The "ripple" ϵ and the "notch" attenuation α are computed by means of Equations (4-6) and (4-7), respectively. Curves of

ϵ vs. Q and ϵ vs. α with N as a parameter for three assumed forms for f(s) are given in Figures 4.2 through 4.7.

a. Assumed Form of Low Pass Transfer Funtion f(s)

Let us assume that:

(1) f(s) is a rational function of the complex (frequency) variable s.

(2) There exists a real value C_1 and a positive integer q such that for large values of $|s|$, f(s) is asymptotic to C_1/s^q.

(3) The poles of f(s) lie in the left-half portion of the complex s plane.

(4) f(o) = 1.

Besides these technical assumptions, let us also assume that the low pass magnitude function $|f(j\omega)|$ is one of the low pass types with a cutoff frequency, $\omega = 1$.

b. Properties of Band Pass Transfer Function g(y)

The band pass transfer function with unity center frequency and percentage bandwidth 1/Q is defined by

$$g(y) = f(jQ(y - 1/y)) \tag{4-1}$$

which has real and imaginary parts Rg(y) and Ig(y), respectively.

From the four assumed properties of f(s) and from Equation (4-1), the following three properties of g(y) can be derived:

(5) $g(y) = g^*(1/y)$ from which it follows that $g'(y) = -(1/y)^2 g^{*\prime}(1/y)$ $= -(1/y)^2 g'^*(1/y)$. (* denotes complex conjugate).

(6) $g(1) = 1$ implying that $Rg(1) = 1$, $Ig(1) = 0$.

(7) $Rg'(1) = 0$.

Proof of (5), (6), and (7):

Since f is a rational function, it can be split into even and odd parts which correspond to the real and imaginary parts of $f(j\omega)$, respectively. Hence, $f^*(-j\omega) = f(j\omega)$ and $g^*(1/y) = f^*[jQ(1/y - y)] = f[-jQ(1/y - y)]$ $= f[jQ(y - 1/y)] = g(y)$, which establishes Property (5).

Property (4) implies Property (6) since $g(1) = g(0) = 1$.

The real part of $f(j\omega)$ is even, implying that its derivative is odd and $\mathrm{Re}[f'(0)] = 0$. Property (7) follows since $\mathrm{Re}[f'(0)] = Rg'(1)$.

$g(y)$ is a band pass filter transfer function with band center frequency $y = 1$. However, the argument of g can be scaled to produce a filter function with a band center at y_o, i.e., $g(y/y_o)$. A set of congruent, evenly-spaced filter functions is given by

$$h_k(x) = g(x/p^k) \text{ for } k = 0, 1, 2, 3, \ldots; -1, -2, -3, \ldots; \tag{4-2}$$

where $p > 1$.

If there are to be N filters per octave, $p = 2^{1/N}$.

c. Properties of $h_k(x)$

We can next establish the following properties of $h_k(x)$:

(8) $h_k(x) = h^*_{-k}(1/x)$.

(9) $h_k(p^k) = 1$.

(10) For m and n nonequal integers, $h_m(p^{(m+n)/2}) + h_n(p^{(m+n)/2})$

$= 2Rg(p^{(m-n)/2})$.

(11) $h'_m(p^{(m+n)/2}) + h'_n(p^{(m+n)/2}) = j(2/p^m)Ig'(p^{(-m+n)/2})$.

(12) $\sum_{k=-\infty}^{\infty} h_k(x)$ converges for all real x.

(13) $\frac{d}{dx}\left[\sum_{k=-\infty}^{\infty} h_k(x)\right] = \sum_{k=-\infty}^{\infty} h'_k(x)$.

Proof of Properties (8) through (13):

$h_k(x) = g(x/p^k) = g*(p^k/x) = g*[(1/x)p^{-k}] = h*_{-k}(1/x)$ establishes Property (8).

$h_k(p^k) = g(p^k/p^k) = g(1) = 1$, which establishes Property (9).

$h_m(p^{(m+n)/2}) + h_n(p^{(m+n)/2}) = g(p^{-(m-n)/2}) + g(p^{(m-n)/2}) = g*(p^{(m-n)/2})$

$+ g(p^{(m-n)/2}) = 2Rg(p^{(m-n)/2})$, which establishes Property (10).

$h'_m(x) + h'_n(x) = \frac{d}{dx}[g(x/p^m) + g(x/p^n)] = (1/p^m)g'(x/p^m) + (1/p^n)g'(x/p^n)$

$= (1/p^m)g'(x/p^m) - (p^n/x^2)g'*(p^n/x)$; hence, $h'_m(p^{(m+n)/2}) + h'_n(p^{(m+n)/2})$

$= (1/p^m)g'(p^{(-m+n)/2}) - (1/p^m)g'*(p^{(-m+n)/2}) = j(2/p^m)Ig'(p^{(-m+n)/2})$,

establishing Property (11).

By comparison to the geometric series, the series of terms $h_k(x)$ converges; that is, for large values of k,

$h_{\pm k}(x) = f[jQ(x/p^{\pm k} - p^{\pm k}/x)] \sim \dfrac{C_1}{[jQ(x/p^{\pm k} - p^{\pm k}/x)]^q} = O[(1/p^q)^k]$, which

establishes Property (12).

By Property (2), f(s), which is analytic in a deleted neighborhood of infinity, can be represented by its Laurent series,

$$f(s) = C_1/s^q + C_2/s^{q+1} + \ldots,$$

and its derivative can be represented by

$$f'(s) = -qC_1/s^{q+1} - (q+1)C_2/s^{q+2} + \ldots$$

Consequently, if x is restricted to a finite closed interval not including zero, if can be readily shown that $h'_{\pm k}(x) = O[(1/p^q)^k]$ for x in the interval, and by the Weierstrass M-test, the series $\sum_{k=-\infty}^{\infty} h'_k(x)$ converges uniformly for all x in the interval. Hence, by a theorem on the differentiability of an infinite series[30] the series can be differentiated term by term at any finite nonzero point, establishing Property (13).

d. Properties of Composite Transfer Response

The "all-pass" transfer response of the infinite filter set is

$$H(\omega) = \sum_{k=-\infty}^{\infty} h_k(\omega) \tag{4-3}$$

Some properties of $H(\omega)$ are:

(14) $H(\omega)$ is log periodic, i.e. $H(\omega) = H(p\omega)$.

(15) $H(\omega) = H^*(1/\omega)$.

(16) $$MH(1) = |H(1)| = 1 + 2 \sum_{k=1}^{\infty} Rg(p^k). \tag{4-4}$$

(17) $$MH(p^{1/2}) = 2 \sum_{k=0}^{\infty} Rg(p^{(k+1)/2}) . \tag{4-5}$$

(18) $MH'(1) = 0.$

(19) $MH'(p^{1/2}) = 0.$ (i.e., relative maxima or minima occur at $\omega = 1$ and $\omega = p^{1/2}$)

Proof of (14): $H(p\omega) = \sum_{-\infty}^{\infty} g(p\omega/p^k) = \sum_{-\infty}^{\infty} g(\omega/p^{k-1}) = \sum_{-\infty}^{\infty} g(\omega/p^{k'})$

---

[30] Olmstead, J.M. · Real Variables, Appleton-Century-Crafts, New York, 1959, p. 283.

$= H(\omega)$ where $k' = k - 1$.

Proof of (15): $H^*(1/\omega) = \sum_{-\infty}^{\infty} h_k^*(1/\omega) = \sum_{-\infty}^{\infty} h_{-k}(\omega) = \sum_{-\infty}^{\infty} h_k(\omega) = H(\omega)$

Proof of (16): $H(1) = \sum_{-\infty}^{\infty} h_k(1) = h_o(1) + \sum_{1}^{\infty} (h_k(1) + h_{-k}(1))$.

Property (10) for $m=k$, $n=-k$ yields $h_k(1) + h_{-k}(1) = 2Rg(p^k)$.

Since $H(1)$ is a real number, $MH(1) = H(1)$.

Proof of (17): $H(p^{1/2}) = \sum_{-\infty}^{\infty} h_k(p^{1/2}) = \sum_{1}^{\infty} (h_{k+1}(p^{1/2}) + h_{-k}(p^{1/2}))$.

Property (10) for $m = k+1$, $n = -k$ yields $h_{k+1}(p^{1/2}) + h_{-k}(p^{1/2})$

$= 2Rg(p^{(2k+1)/2})$, a real number.

Proof of (18): $MH^2(\omega) = |H(\omega)|^2 = Re^2[H(\omega)] + Im^2[H(\omega)]$.

The derivative of MH can be computed as

$$MH'(\omega) = \frac{1}{MH(\omega)} \left[Re[H(\omega)] \,.\, Re[H'(\omega)] + Im[H(\omega) \,.\, Im[H'(\omega)] \right]$$

Then $MH'(1) = \frac{1}{MH(1)} \left[H(1).Re[H'(1)] + 0 \right]$.

$$H'(1) = \sum_{-\infty}^{\infty} h_k'(1) = g'(1) + \sum_{k=1}^{\infty} (h_k'(1) + h_{-k}'(1))$$

$= g'(1) + j \sum_{k=1}^{\infty} (2/p^k) Ig'(p^{-k})$, by Property (11) with $m=k$ and $n=-k$, and

$Re[H'(1)] = Rg'(1) = 0$.

Proof of (19): $MH'(p^{1/2}) = \frac{1}{MH(p^{1/2})}\left[H(p^{1/2}).Re[H'(p^{1/2})] + 0\right]$

$$H'(p^{1/2}) = \sum_{0}^{\infty} \left[h_{k+1}(p^{1/2}) + h_{-k}(p^{1/2})\right]$$

$= j \sum_{0}^{\infty} (2/p^{k+1}) Ig'[p^{(-2k+1)/2}]$, the real part of which is zero.

e. Ripple and Notch Calculations for Three Assumed Forms of f(s)

As was indicated previously, the "ripple" ϵ, which is a measure of the deviation of the frequency response from a flat characteristic, is defined as

$$\epsilon = 20 \log_{10} \frac{Max[|H(\omega)|]}{Min[|H(\omega)|]} \tag{4-6}$$

and the "notch" attenuation is defined by

$$\alpha = 20 \log_{10} \frac{H(1)}{[H(1) - h_o(1)]} \tag{4-7}$$

When Q is sufficiently large, we can make the approximations,

$$Max[|H(\omega)|] \doteq H(1) = 1 + 2 \sum_{k=1}^{\infty} Rg(p^k), \tag{4-8a}$$

$$Min[|H(\omega)|] \doteq H(p^{1/2}) = 2 \sum_{k=0}^{\infty} Rg(p^{k+1/2}). \tag{4-8b}$$

Therefore, by substituting Equations (4-8a) and (4-8b) into Equations (4-6) and (4-7) we can derive the following formulas for ϵ and α:

$$\epsilon \doteq 20 \log_{10}\left\{\frac{1 + 2 \sum\limits_{k=1}^{\infty} Rg(p^k)}{2 \sum\limits_{k=0}^{\infty} Rg(p^{k + 1/2})}\right\} \text{ decibels.} \qquad (4\text{-}9)$$

$$\alpha = 20 \log_{10}\left[1 + \frac{1}{2 \sum\limits_{k=1}^{\infty} Rg(p^k)}\right] \text{ decibels.} \qquad (4\text{-}10)$$

Except for the first-order filter, if Q is not sufficiently large, the true value of ϵ will be larger than that given by Equation (4-9) since the true maximum and minimum values of $|H(\omega)|$ are liable to occur at other points than $\omega = 1$ and $\omega = p^{1/2}$. Under these conditions we only know that ϵ cannot be less than the value given by this equation. However, correct values of ϵ for small values of Q will be computed by a procedure outlined in Section 4.4.

Values of ϵ and α given by Equations (4-9) and (4-10) were calculated using the IBM 7094 computer for various values of N and Q and for three forms of f(s). These are the Butterworth functions of orders one, two, and three. The magnitudes of these functions for $s = j\omega$ are:

$$Mf_1(j\omega) = \frac{1}{\sqrt{1 + \omega^2}}, \qquad (4\text{-}11a)$$

$$Mf_2(j\omega) = \frac{1}{\sqrt{1 + \omega^4}}, \qquad (4\text{-}11b)$$

$$Mf_3(j\omega) = \frac{1}{\sqrt{1 + \omega^6}}, \qquad (4\text{-}11c)$$

These correspond to the following complex filter functions:

$$f_1(s) = \frac{1}{1 + s}, \qquad (4\text{-}12a)$$

$$f_2(s) = \frac{1}{1 + \sqrt{2}s + s^2}, \tag{4-12b}$$

$$f_3(s) = \frac{1}{1 + 2s + 2s^2 + s^3}, \tag{4-12c}$$

whose real parts for $s = j\omega$ are given by

$$Rf_1(\omega) = \frac{1}{1 + \omega^2}, \tag{4-13a}$$

$$Rf_2(\omega) = \frac{1 - \omega^2}{1 + \omega^4}, \tag{4-13b}$$

$$Rf_3(\omega) = \frac{1 - 2\omega^2}{1 + \omega^6}, \tag{4-13c}$$

f. Numerical Evaluation of the Infinite Series

Given the real part of $f(j\omega)$ and values of Q and $N = \log_2(p)$, the two series to be evaluated are

$$\sum_{k=1}^{\infty} Rf[Q(p^k - p^{-k})] \quad \text{and} \quad \sum_{k=0}^{\infty} Rf[Q(p^{(k+1)/2} - p^{(-k-1)/2})]$$

It will be clear from the discussion that follows that a finite sum approximation of the second series will be at least as accurate as that of the first series for the same number of terms. Therefore, only the first sum will be treated.

Suppose that, for k greater or equal to some fixed integer K, the general term of a series a_k satisfies the condition

$$a_{k+1} \le r\, a_k \quad \text{where } r < 1. \tag{4-14}$$

It follows that, for i, a non-negative integer,

$$a_{k+i} \le a_k r^i. \tag{4-15}$$

The error between the value of the infinite series S and the finite approximation to S, S_k, is

$$\text{Error} = \left|S - S_k\right| = \left|\sum_{i=k+1}^{\infty} a_i\right| \le \sum_{i=k+1}^{\infty} \left|a_i\right|$$

$$\le \sum_{i=0}^{\infty} \left|a_k\right| r^i - \left|a_k\right| = \left|a_k\right| (1/(1-r)-1),$$

which reduces to

$$\text{Error} \le \frac{\left|a_k\right|}{(1/r-1)}, \quad k \ge K. \tag{4-16}$$

Then the fractional error is restricted by

$$\text{F.E.} \le \frac{\left|a_k\right|}{S_k} \frac{1}{(1/r - 1)}, \quad k \ge K. \tag{4-17}$$

Given that $\underline{\omega}(k) = Q(p^k - p^{-k}) = 2Q\sinh[k\cdot\ln(p)]$ and $a_k = Rf[\underline{\omega}(k)]$, Condition (4-14) can be rewritten as

$$1 < \frac{1}{r} \le \frac{1/\left|Rf[\underline{\omega}(k+1)]\right|}{1/\left|Rf[\underline{\omega}(k)]\right|}, \quad k \ge K. \tag{4-18}$$

which is equivalent to

$$0 < \ln(\frac{1}{r}) \le \ln\left[\frac{1}{\left|Rf[\underline{\omega}(k+1)]\right|}\right] - \ln\left[\frac{1}{\left|Rf[\underline{\omega}(k)]\right|}\right]. \tag{4-19}$$

By the Law of the Mean there exists a number z between k and k + 1 such that

$$\frac{d}{dx}\left[\ln\left(\frac{1}{|Rf[\underline{\omega}(x)]|}\right)\right]_{x=z} = \ln\left[\frac{1}{|Rf[\underline{\omega}(k+1)]|}\right] - \ln\left[\frac{1}{|Rf[\underline{\omega}(k)]|}\right]. \qquad (4\text{-}20)$$

Hence, a condition which is sufficient to imply Condition (4-18) for $x \geq K$ is

$$\frac{-Rf'[\underline{\omega}(x)]\ \underline{\omega}'(x)}{Rf[\underline{\omega}(x)]} \geq t > 0. \qquad (4\text{-}21)$$

Note that

$$\underline{\omega}'(x) = 2Q\ \ln(p)\ \cosh\ [(x\ \ln(p)] \geq 2Q\ \ln(p)\ \sinh\ [x\ \ln(p)], \qquad (4\text{-}22a)$$

or

$$\underline{\omega}'(x) \geq \underline{\omega}(x)\ln(p). \qquad (4\text{-}22b)$$

Finally, since $\underline{\omega}$ is a monotonically increasing function, Condition (4-18) can be replaced by the criterion

$$\frac{-Rf'(\omega)\ \omega\ \ln(p)}{Rf(\omega)} \geq t > 0, \quad \omega \geq W > 0. \qquad (4\text{-}23)$$

The constants need for the evaluation of the fractional error by Equation (4-17) are given by

$$r = e^{-t} \qquad (4\text{-}24a)$$

$$k \geq \frac{NW}{2Q\ \ln(2)} > \frac{N}{\ln(2)}\ \sinh^{-1}\left[\frac{W}{2Q}\right] = \underline{\omega}^{-1}(W). \qquad (4\text{-}24b)$$

An analysis of the real parts of the first three Butterworth low pass filter functions given by Equations (4-13a,b,c), indicate that adequate values of r and K for the three cases are·

$$r_1 = r_2 = 1/p, \ r_3 = 1/p^2,$$

$$K_1 = K_3 = N/Q, \ K_2 = 1.4N/Q$$

Values of ripple ϵ and notch attenuation α defined by Equations (4-9) and (4-10) were next computed such that the accuracy of the infinite sums involved were within 1 per cent of their correct values according to Criterion (4-17). Graphs of ϵ and α in decibels are given in Figures 4.2 through 4.7 for values of Q between 1 and 20 and N between 1 and 10 for the three filter types. Note that in each case for large values of N, the curves of α vs. Q (Figures 4.3, 4.5, 4.7) tend to converge to limiting positions.

4.4 Calculations for Finite Sets of Band Pass Filters

In the case of the infinite filter set utilizing first-order filters, there is only one pole per interval of repetition of the transfer function. Therefore there can only be one maximum and one minimum value of the transfer function per interval. From the analysis given above, these critical values must occur at the band pass filter center frequencies and the geometric mean frequencies between them. In this case, the absolute value of ϵ gives the correct ripple value for all values of Q and N. However, for higher order filters the number of poles per interval is greater than one, and the possibility exists, for comparatively large values of N/Q, that other maximums and minimums will occur. Consequently, a digital computer program utilizing complex number arithmetic, was written to calculate several points per repetition interval of the composite transfer function of n filters in order to determine the values of true ripple for comparatively large values of N/Q. That is, the function

$$\sum_{k=0}^{n-1} f[jQ(\omega/p^k - p^k/\omega)] \qquad (4\text{-}25)$$

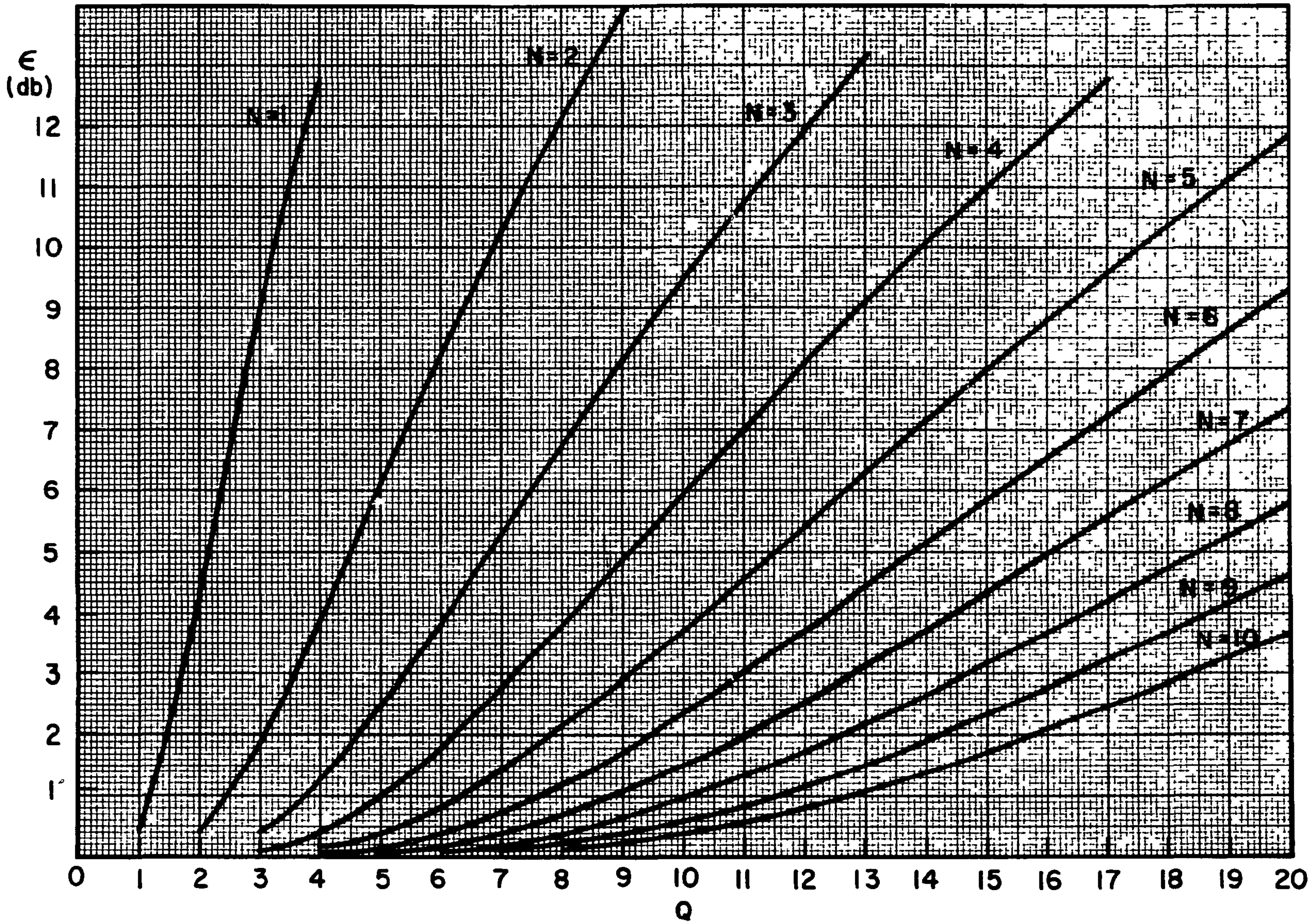


Figure 4.2 Infinite Filter Set: Ripple ε vs. Filter Q (Filter Density N is a Parameter), First Order Butterworth.

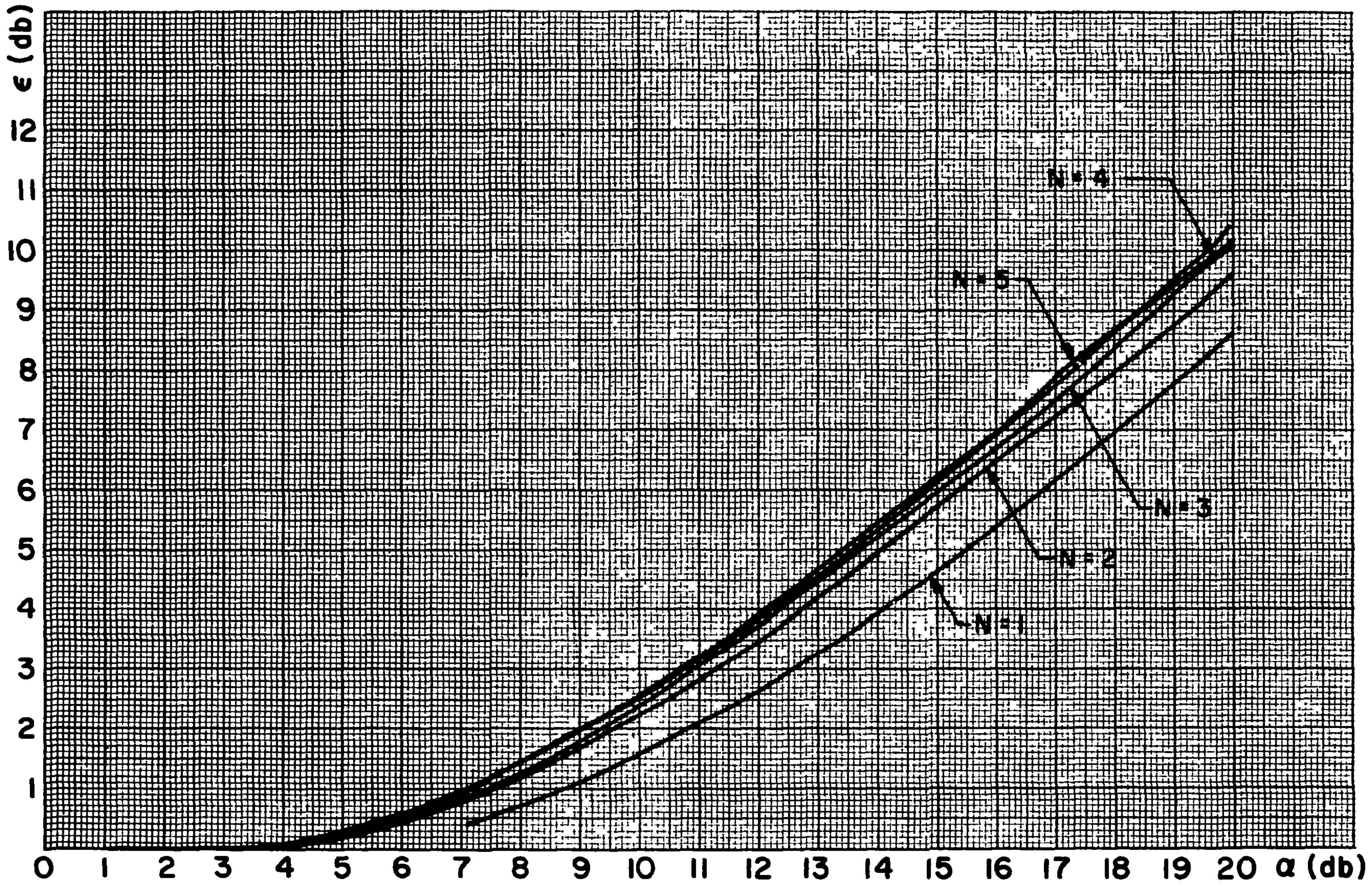


Figure 4.3 Infinite Filter Set: Ripple ε vs. Notch Attenuation α (Filter Density N is a Parameter), First Order Butterworth.

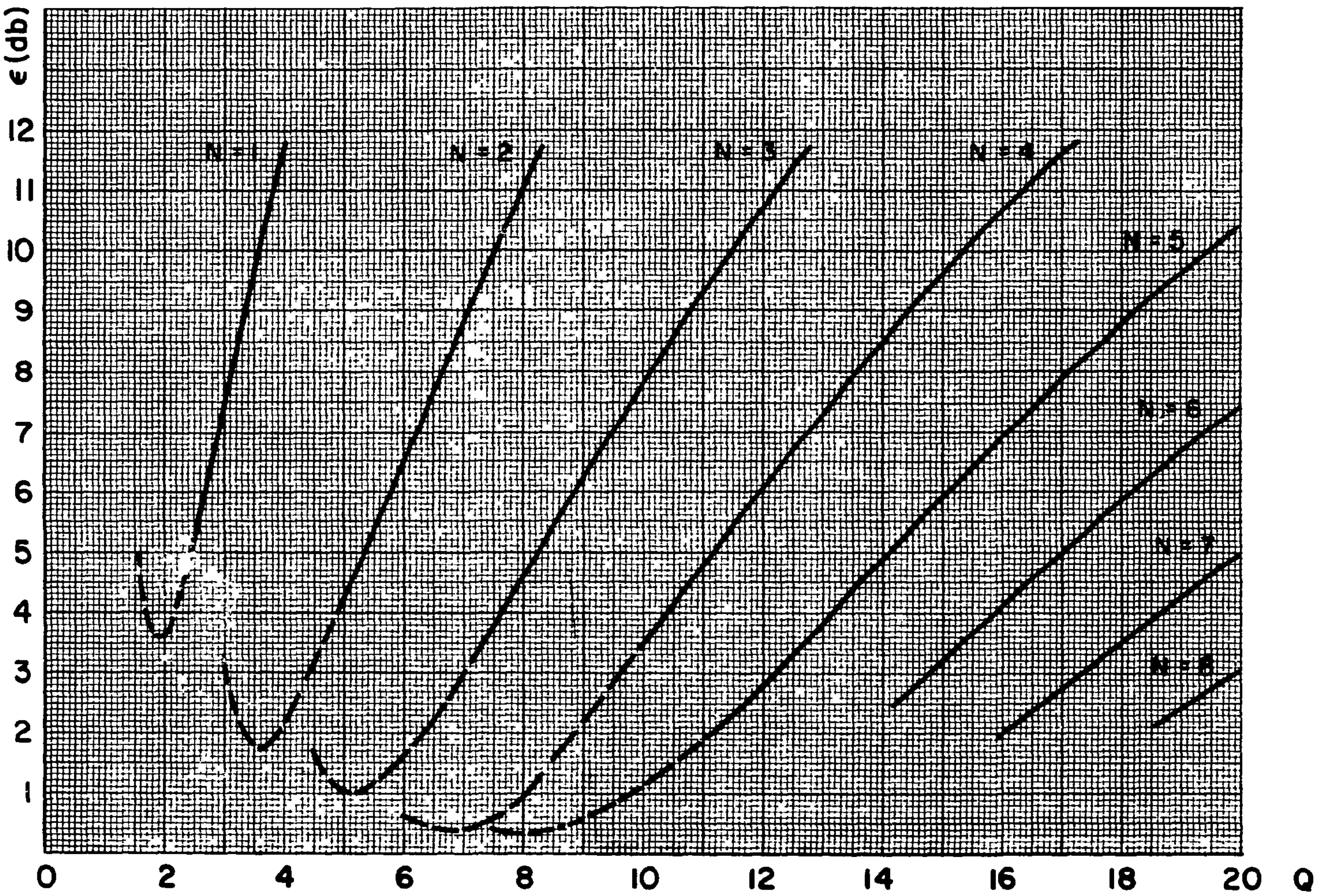


Figure 4.4 Infinite Filter Set: Ripple ε vs. Filter Q (Filter Density N is a Parameter), Second Order Butterworth.

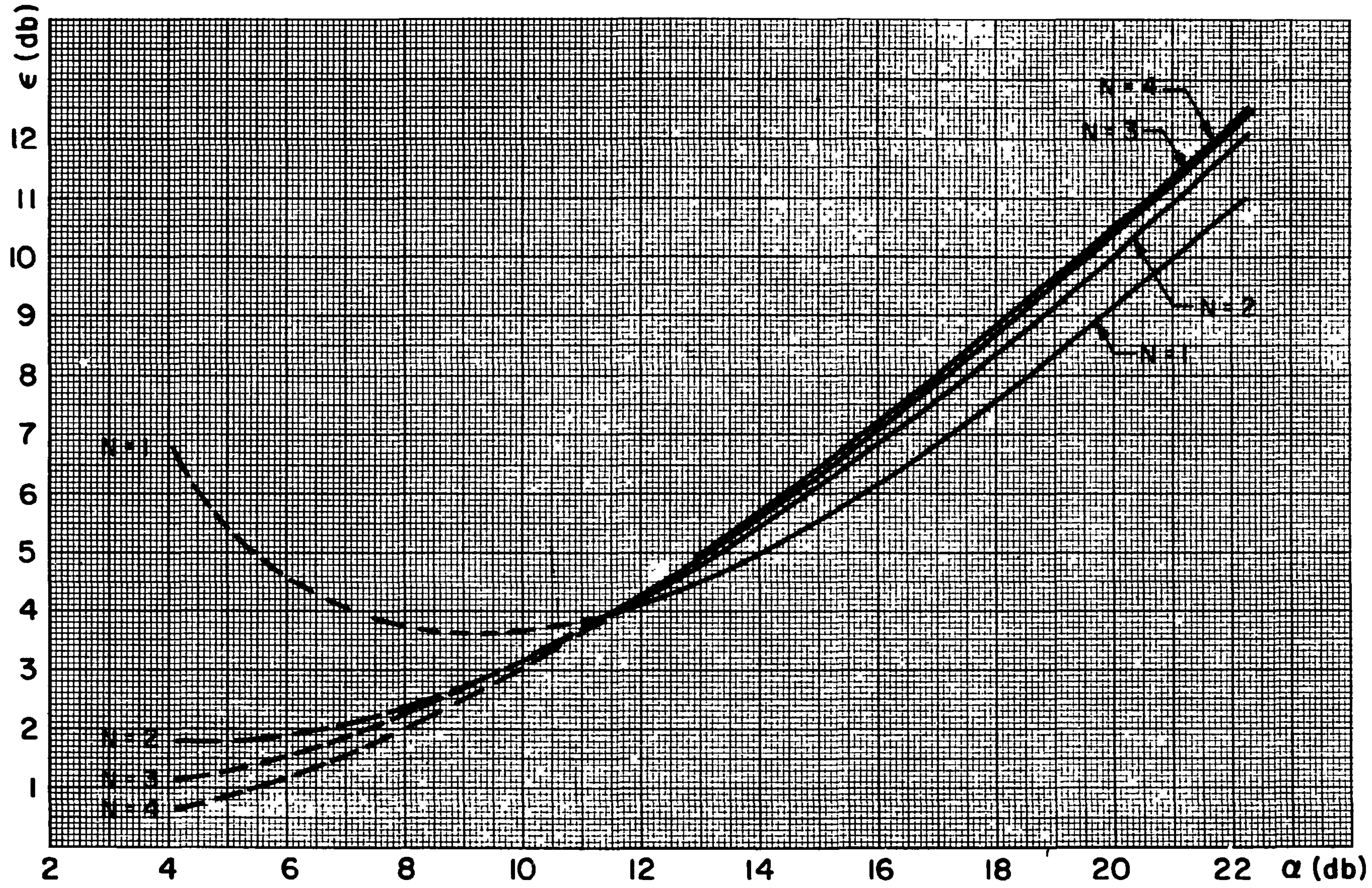


Figure 4.5 Infinite Filter Set: Ripple ϵ vs. Notch Attenuation α (Filter Density N is a Parameter), Second Order Butterworth.

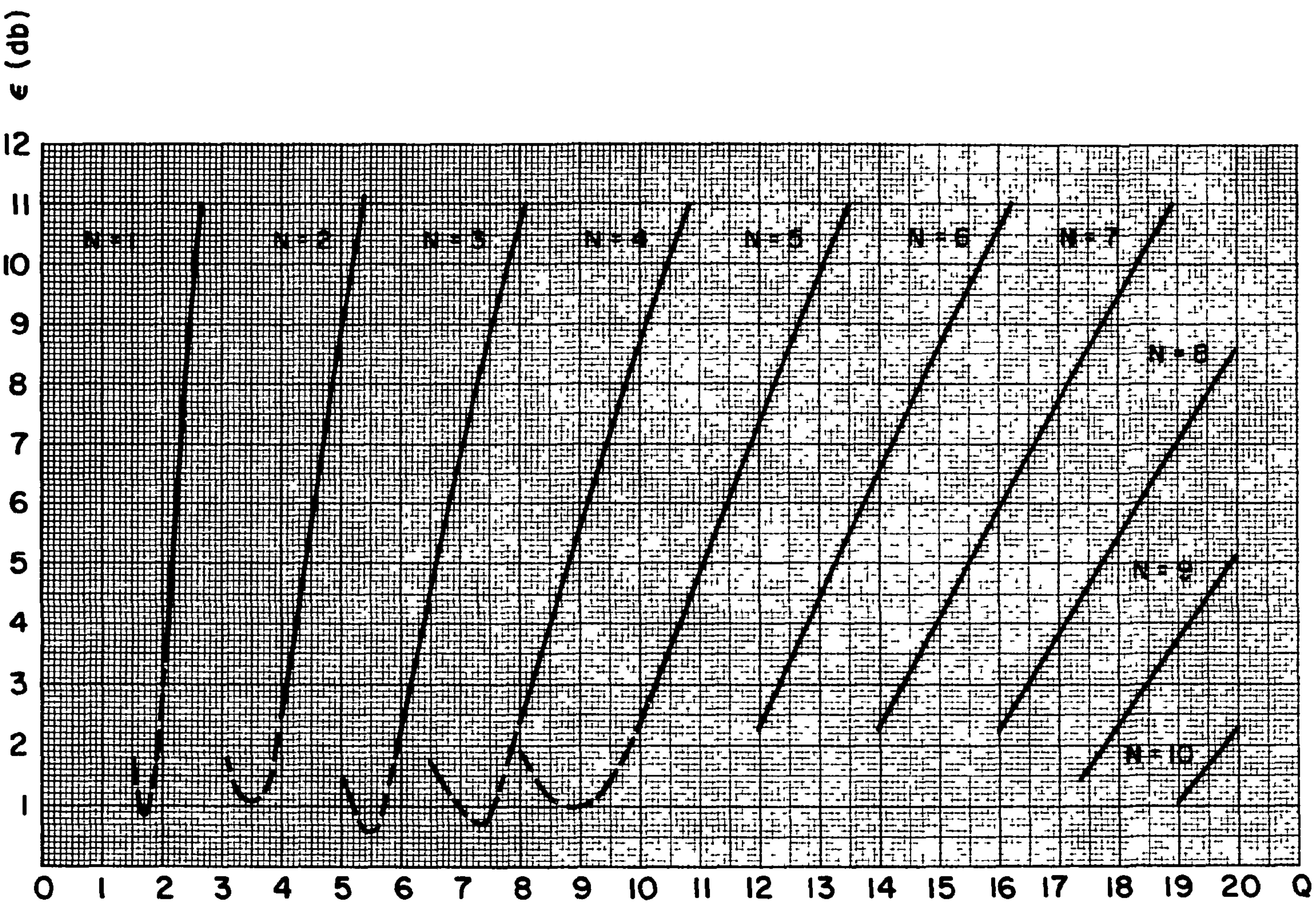


Figure 4.6 Infinite Filter Set: Ripple ε vs. Filter Q (Filter Density N is a Parameter), Third Order Butterworth.

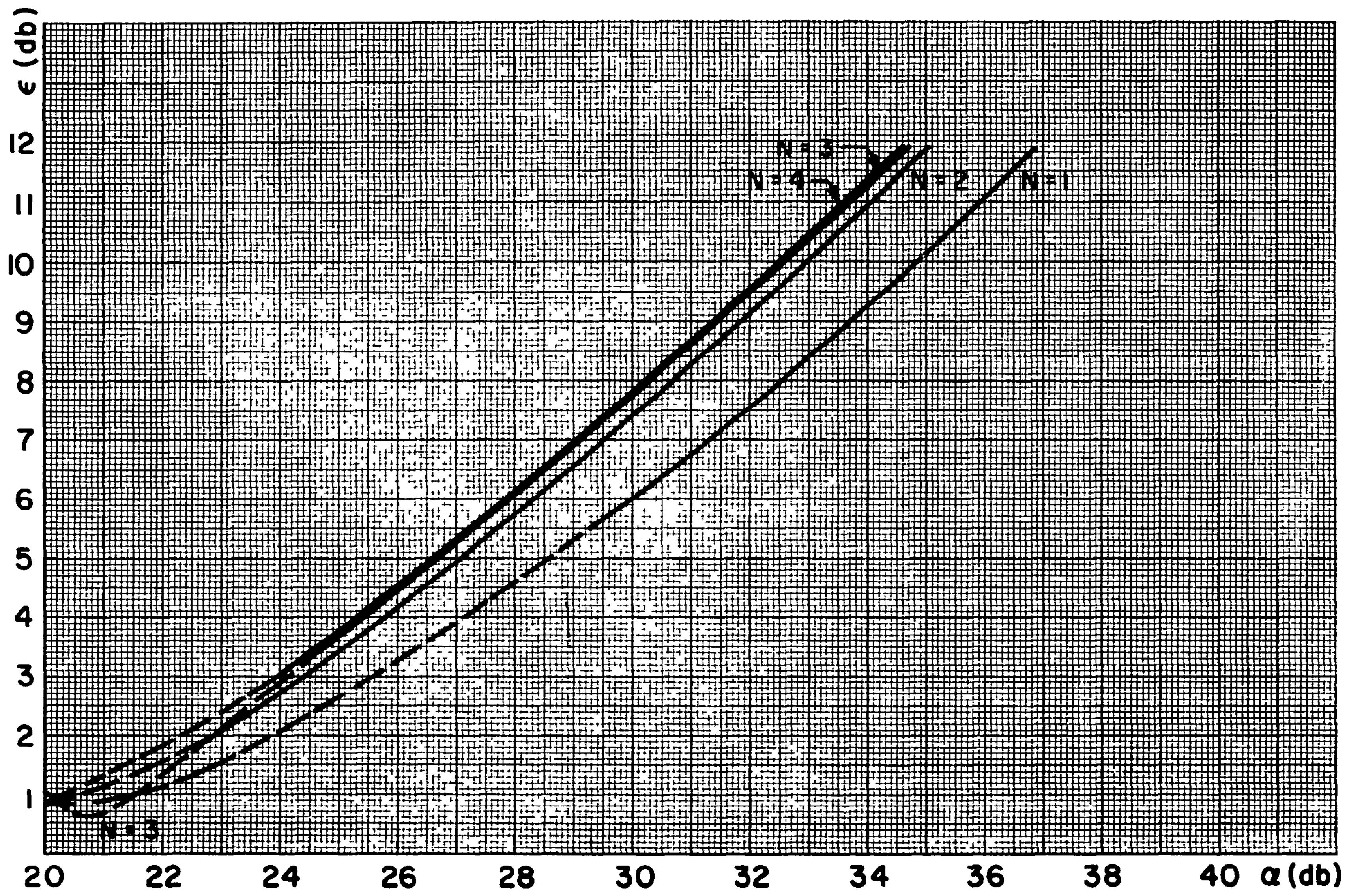


Figure 4.7 Infinite Filter Set: Ripple ε vs. Notch Attenuation α (Filter Density N is a Parameter), Third Order Butterworth.

was plotted for values of ω between 1/2 and 2 p^{n-1} for appropriate values of n, N(where $p = 2^{1/N}$) and Q for the second and third-order cases. The precision of the values derived in this manner, as applied to infinite filter sets of the same type, was inferred from the periodicity of the transfer function, which is optimum on the interval midway between $\omega = 1$ and $\omega = p^{n-1}$. The curves of ϵ vs. Q and α vs. ϵ were actually derived by adjoining the data obtained by the infinite filter set analysis (solid curves) with those obtained from the analysis of sets with a large number of filters (dashed curves).

4.5 A Practical Variable Formant Filter

Because of cost considerations, the breadth of the conventional rack mount panel, and the width of available linear attenuators, the number of filters was restricted to 12 for the first prototype model constructed. Also, due to the cost of manufacturing higher order band pass filters (utilizing a corresponding larger number of components with increasingly stringent tolerances), the order of the filters under consideration was limited to three. A decision was made, based on subjective experience and from the review of available data on musical instrument formants,[2] that a filter density of 3 filters per octave would satisfy the requirements for selectivity and that the frequency range 189 cps to 2400 cps would cover the principal region of interest for musical formants. Also the ripple was specified not to exceed 3 db and the notch attenuation due to the infinite attenuation of a single mid-band filter was to be at least 15 db. According to the data represented by the curves given, the largest values of α obtainable under these conditions are 10.8, 10.1, and 22.9 db for the 1st, 2nd, and 3rd order Butterworth filters, respectively. Note that the second-order filters offer no improvement over the first order types. It was clear that the third-order types were required. Since the minimum ripple possible for this type for N = 3 is .6 db (a critical point where equi-ripple response occurs), corresponding to $\alpha =$ 20.6 db (only slightly worse than 22.9 db), filters which achieve this characteristic having Q = 5.5 were chosen. The specifications for the band pass filters for the Variable Formant Filter are:

1. Response Shape: Maximally Flat (Butterworth) Third Order.
2. Percentage Bandwidth· 18 per cent.
3. Center Frequencies· 189, 238, 300, 378, 475, 600, 756, 953, 1200 1510, 1905, and 2400 cps.

The response characteristics of filters supplied by White Instrument Laboratories (Austin, Texas) were measured. The center frequencies were all within 3 per cent of the specified values, the percentage bandwidth varied from 17.2 percent to 18.3 percent and the attenuations at twice the center frequency were all within 4 db of 55 db, the predicted attenuation.

The ripple of the summed response of the Variable Formant Filter, which was constructed using these band pass filters, was measured to be less than 1 db and the notch attenuation due to the removal of a single filter at midband was measured to be approximately 21 db.

Variable linear attenuators (Lumiten 678) were purchased from Fairchild Recording Equipment Corporation. The vertical position of a knob mounted on the front panel of one of these units corresponds to the attenuation of an input signal, as is indicated by a decibel scale drawn on the face of the front panel. The knob varies the brightness of a lamp by means of a variable resistor in series with a fixed voltage (6.3 v.). The lamp, in turn, operates to vary the resistance of a light-sensitive photoresistor, which is the attenuating element of the circuit.

There are two principal advantages with respect to using this type of attenuator in musical applications. First, the fact that the attenuation of each filter output is indicated vertically means that formant contours set up visually according to the positions of the knobs will at least in a qualitive sense correspond to the actual formant characteristic. Second, the use of a lamp and a light-controlled resistor to smooth out (by virtue of the sluggishness of their responses) the abrupt changes between resistor contacts prevents the introduction of "noise" into the signal output, which ordinarily occurs when the control knob is moved rapidly across a wire-wound resistor.

Unfortunately, in the process of constructing the Variable Formant Filter, we also discovered several disadvantages with these attenuators. First, the actual measured attenuations did not closely coincide with the scales marked on the front panels because the characteristics of the lamps and photoresistors vary considerably from attenuator to attenuator. Second, the attenuation is very sensitive to changes in lamp voltage; hence, the lamp voltage must be held within .1 volt. Third, the lamps may burn out, causing total failure of the attenuator. Fourth, the response to an increase in attenuation is quite slow due to the finite time it takes a lamp to lose its brightness. Fifth, there is no infinite attenuation setting, the lowest setting possible is about 70 db.

While experimenting with the Variable Formant Filter, we found that the maximum notch attenuation does not occur when a mid-band attenuator is set for zero gain. The greatest notch attenuation was measured to be 40 db when the filter attenuation was set to 21 db.

The present Variable Formant Filter does not cover the entire useful frequency range for electronic music applications. Frequencies below 189 cps. and above 2400 cps. are severely attenuated. In the near future, we expect to expand this instrument in order to cover the entire range, 60 to 16,000 cps. Since the ear is not as selective in separating frequency bands in the lower and upper extremes of the audio spectrum, it may be sufficient to use filter sets with lower filter densities in these regions. Alternatively, filters with sharp-cutoff characteristics on the sides adjacent to the rest of the filter set and with variable "roll-off" characteristics on the other sides of their center frequencies might be used to fill out the high and low ends of the audio frequency spectrum.

4.6 Computer Simulation of Variable Formant Filters

The same program used to evaluate the ripple of a finite filter set can be used to simulate an adjustable formant filter. That is, the frequency response of this filter, specified by the filter density N, the total number of filters n, the individual complex filter transfer function [given by the low pass equivalent, f (s)], and the filter Q, is printed out for a specified set of filter attenuations in decibels. Figure 4.8 illustrates the results of an attempt to fit the decibel curve $-20 \log_{10}\omega$ with the frequency response of the variable filter which was physically realized ($N=3$, $n=12$, $f(s) = 1/(s^3 + 2s^2 + 2s + 1)$, $Q = 5.5$). The filter attenuations for this case are given by $\alpha_k = 2(k-1)$ db. The graph is actually a plot of the equation

$$H(\omega) = 20 \log_{10} \left| \sum_{k=0}^{n-1} 10^{(k-1)/10} f\left(jQ[\omega/2^{k/N} - 2^{k/N}/\omega]\right) \right|, \qquad (4\text{-}26)$$

for $.5 \leq \omega \leq 20$.

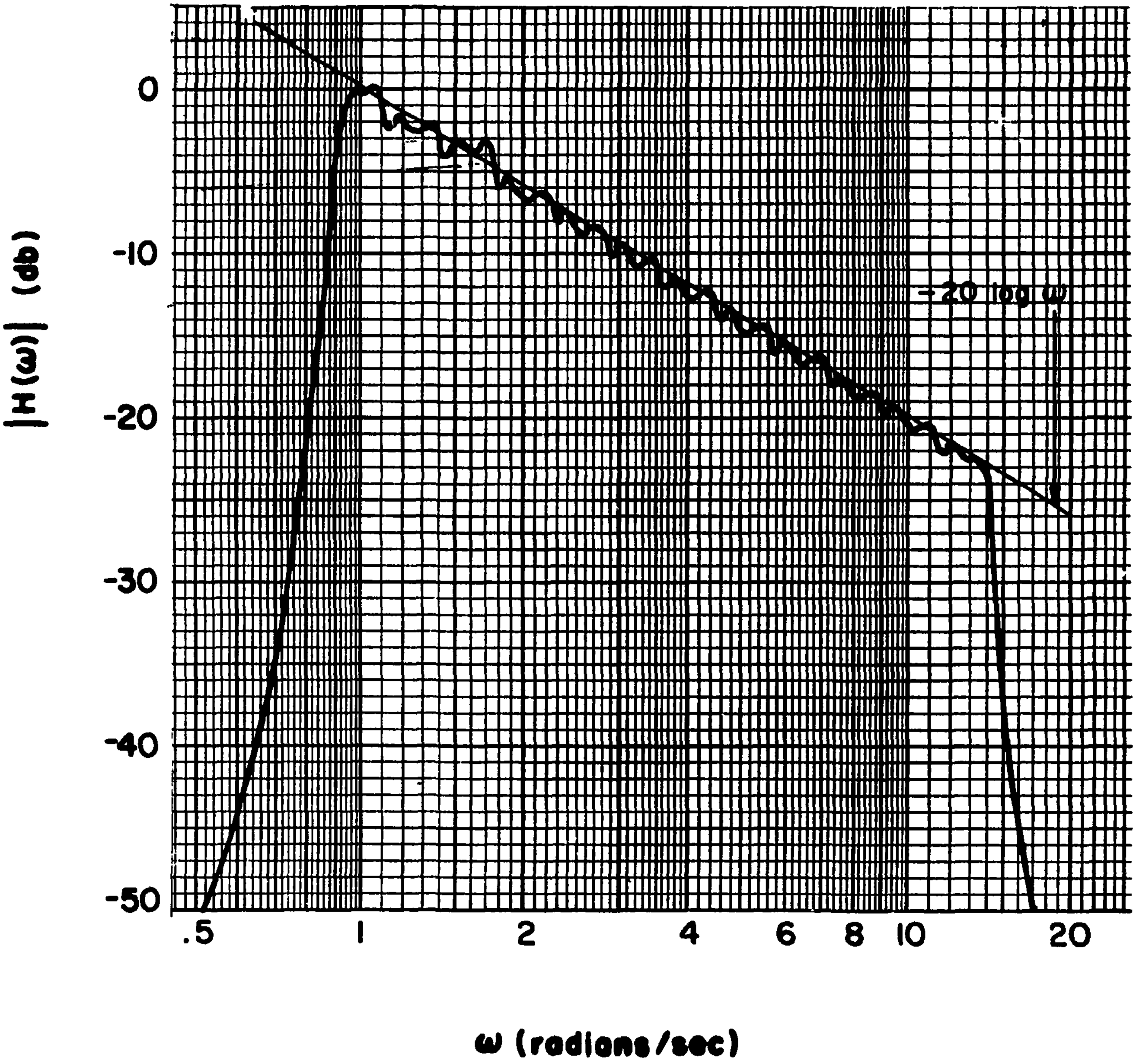


Figure 4.8 Variable Formant Filter Simulation of $-20\log_{10}(\omega)$.

Chapter 5

FOURIER SERIES METHODS FOR ANALYSIS OF A TRANSIENT HARMONIC TONE AND A PROJECTED ANALOG SYSTEM FOR ANALYSIS AND RESYNTHESIS OF MUSICAL TONES

5.1 Introduction

In this chapter, two "complete" methods for representing a transient harmonic signal by a time-variant Fourier series are derived. In both cases the computation of the Fourier coefficients, which vary with time, can be divided into two steps: (1) Hetrodyning the signal with quadrature pairs of each harmonic of the signal fundamental frequency, and (2) averaging or low-pass filtering of the hetrodyned components. Although Step (1) is unique, there is a considerable amount of freedom for specifying Step (2), the low pass filter operation. In Section 5.6, a third low pass operation is found to yield a more accurate analysis of the test signal $U(t)e^{-t/\tau}\cos(k2\pi t)$ (where k is an integer) than the other two types which correspond to "complete" methods for analysis.

Finally, a plan for an electronic device which could be used to analyze musical sounds and which is an analog of the general method of time-variant Fourier series calculation, is presented. This system makes use of the ultrasonic frequency techniques developed in designing the Harmonic Tone Generator. The outputs of the analyzer could be used as inputs to an expanded version of the Harmonic Tone Generator which would then resynthesize the analyzed signal.

As we have already seen, a harmonic tone generator can be used to generate a large class of signals. A subclass of possible signals of this type would be classified by musical listeners as consisting of harmonic tones even though these tones are, strictly speaking, nonperiodic. It is only important that the perception of harmonicity be the predominant element which determines the timbre of a tone. A more analytic designation for this type of tone would be "quasi-periodic signal". Deviations from periodicity can be caused by changing fundamental frequency or by modulations of the harmonic amplitudes and phases. [Note that phase time-variations according to the relation $\phi_k(t) = \theta_k + k[\phi_1(t)-\theta_1]$ are most conveniently interpreted in terms of the equivalent fundamental frequency variation, $f = \phi_1'(t)$] .

Besides the problem of generating synthetic harmonic tones with time-varying parameters, it is equally important to consider the problem of determining the parameter specifications for signals which originate as acoustical events and which are known to be substantially harmonic, i.e., quasi-periodic. These signals include the tones of most musical instruments, certain sounds of nature, and the vowel sounds of human speech. Actually,there is one problem with speech sounds as compared to many musical sounds, which is the fact that their analysis in terms of harmonic structure is made difficult by the highly transient behavior of the fundamental frequency. It is much easier to measure the movements of vowel formant peaks by means of time-variant Fourier transform or parallel-filter analyzer systems. Usually, information about the phases of the frequency components is not retained, and the analysis method has no guarantee of "completeness". However, since, in the case of musical tones, pitch is a relatively constant parameter, it is feasible to take advantage of their harmonic character and analyze by time-variant Fourier series methods instead of the more traditional time-variant Fourier transform or parallel filter approaches.

5.2 Fourier Series Representation of a Quasi-Periodic Signal

The well-known theorem of Fourier, which allows the computation of the harmonic structure of a periodic signal from the wave-shape of one of its cycles, can be extended to cover the cases of nonperiodic and quasi-periodic functions. For the case of a function g(t) which is periodic with period T (i.e., $g(t + T) = g(t)$) and satisfies certain other conditions,[31] the coefficients of the Fourier series are given by the following formulas:

$$a_o = \frac{1}{T}\int_{t' - T/2}^{t' + T/2} g(x)dx, \quad a_k = \frac{2}{T}\int_{t' - T/2}^{t' + T/2} g(x) \cos\left[\frac{2\pi kx}{T}\right]dx$$

$$b_k = \frac{2}{T}\int_{t' - T/2}^{t' + T/2} g(x) \sin\left[\frac{2\pi kx}{T}\right]dx, \quad k = 1,2,3,\ldots \tag{5-1}$$

where t' is an arbitrary time value. Then, the function can be expressed for all time by the infinite Fourier series in the following manner:

$$g(t) = a_o + \sum_{k=1}^{\infty}\left[a_k \cos\left[\frac{2\pi kt}{T}\right] + b_k \sin\left[\frac{2\pi kt}{T}\right]\right] \tag{5-2a}$$

$$= a_o + \sum_{k=1}^{\infty} c_k \cos\left[\frac{2\pi kt}{T} - \phi_k\right] \tag{6-2b}$$

---

[31] Churchill, R.V.: Fourier Series and Boundary Value Problems, McGraw-Hill, New York, 1941, p.70.

where $c_k = \sqrt{a_k^2 + b_k^2}$

and
$$\phi_k = \begin{cases} \tan^{-1}\left[\frac{b_k}{a_k}\right] + \frac{\pi}{2}\left[\frac{a_k}{|a_k|} - 1\right] & a_k \neq 0; \\ \frac{\pi}{2} \cdot \frac{b_k}{|b_k|}, & a_k = 0,\ b_k \neq 0; \\ 0, & a_k = 0,\ b_k = 0. \end{cases} \tag{5-2c}$$

If g(t) is a nonperiodic signal it can still be represented by a Fourier series in an arbitrary interval (t_1, t_2). It is only necessary to construct another function f(t) such that f(t) = g(t) for $t_1 < t < t_2$ and $f(t + t_2 - t_1) = f(t)$ to see that this is true. That is, f(t) is equal to g(t) in the interval (t_1, t_2) and is the periodic continuation of g(t) outside this interval. Since f(t), a periodic function, has a Fourier series representation good for all time, it follows that g(t) also has a Fourier series representation in the interval (t_1, t_2)--indeed, it is the same series. Mathematically this is expressed by

$$g(t) = f(t) = a_o + \sum_{k=1}^{\infty}\left[a_k \cos\left[\frac{2\pi kt}{T}\right] + b_k \sin\left[\frac{2\pi kt}{T}\right]\right], \quad t_1 < t < t_2 \tag{5-3}$$

where $T = t_2 - t_1$. In particular, for $t = t' = \frac{t_1 + t_2}{2}$,

$$g(t') = a_o + \sum_{k=1}^{\infty}\left[a_k \cos\left[\frac{2\pi kt'}{T}\right] + b_k \sin\left[\frac{2\pi kt'}{T}\right]\right]. \tag{5-4}$$

where the coefficients are given by Equation (5-1). However, it is clear that any interval (t_1,t_2) can be chosen to compute such a Fourier series and we are free to choose a set of intervals such that $t_2 - t_1 = T$ is a constant and $t' = \frac{t_1 + t_2}{2}$ is variable, taking on any value of time. In this case the the coefficients given by Equation (5-1) depend on t' and Equation (5-4) can be rewritten

$$g(t') = a_o(t') + \sum_{k=1}^{\infty} \left[a_k(t') \cos\left(\frac{2\pi kt'}{T}\right) + b_k(t') \sin\left(\frac{2\pi kt'}{T}\right)\right] \qquad (5\text{-}5a)$$

$$= a_o(t') + \sum_{k=1}^{\infty} c_k(t') \cos\left[\frac{2\pi kt'}{T} - \phi_k(t')\right] . \qquad (5\text{-}5b)$$

As long as g(t') is piecewise continuous,[32] the coefficients are continuous functions of time. Further, if the function is initially nonperiodic but becomes periodic with time interval T after a certain period of time, we are guaranteed that the coefficients will become constants. In any case, because of the Fourier theorem, we know that this Fourier series with time-variant coefficients is a "complete" representation of the function g(t); that is to say, g(t) can be reconstructed perfectly from a knowledge of these coefficients.

A modification of Equation (5-1) was used by Luce[33] to compute the instantaneous amplitudes and phases of the individual harmonics of a number of musical instrument tones. He chose the set of intervals (t_1,t_2) to coincide with the successive fundamental periods of the transient harmonic signals being analyzed (i.e., (0,T), (T,2T),...). Therefore, only one computation per period was performed, and since the signals were nonperiodic, the coefficients were bound to change discretely at the end points of each fundamental period. In this case, the representation of a nonperiodic signal given by Equation (5-5) is not strictly correct. For at these end points the Fourier series will converge to $[g(t') + g(t' - T)]/2$ instead of $g(t')$. On the other hand, Luce's results indicate that in many cases the amplitudes and phases follow a smooth curve with respect to time so that a continuous curve might be inferred from it.

---

[32] Reference 31, p.65.

[33] Reference 2, pp. 20-217.

5.3 Frequency Domain Equivalent of Fourier Series Coefficient Calculation

Although Equation (5-5) is a "complete" representation, one might ask the question: How meaningful are the coefficient functions, $a_o(t)$, $a_k(t)$, $b_k(t)$, and their relatives, $c_k(t)$ and $\phi_k(t)$, for a given signal $g(t)$? I propose to sidestep this question by posing a simpler question: Given a function which can be expressed by the equation

$$g(t) = \sum_{k=1}^{n} \bar{c}_k(t) \cos\left[\frac{2\pi kt}{T} + \bar{\phi}_k(t)\right] , \tag{5-6}$$

where $\bar{c}_k(t)$ and $\bar{\phi}_k(t)$ have a restricted amount of change over any given period $(t, t+T)$, what is the relation of the $c_k(t)$'s and $\phi_k(t)$'s computed by Equations (5-1) and (5-2c) to the original $\bar{c}_k(t)$'s and $\bar{\phi}_k(t)$'s? Although no simple answer to this question can be given, it is possible to show by means of a frequency domain equivalent of the time-varying Fourier series calculation method that a certain amount of information contained in an original envelope $\bar{c}_k(t)$ will, because of its transient behavior, be missing in the corresponding computed

$c_k(t)$. Moreover, the graph of each $c_k(t)$ will show a certain amount of ripple due to the transient behavior of any of the parameters including the original $\bar{c}_k(t)$.

Note that Equations (5-1) and (5-5) can be combined to yield the identity:

$$g(t) = \frac{1}{T} \int_{t - T/2}^{t + T/2} g(x)dx + \sum_{k=1}^{\infty} \left\{ \left[\frac{2}{T} \int_{t - T/2}^{t + T/2} g(x) \cos\left(\frac{2\pi kx}{T}\right) dx \right] \cos\left(\frac{2\pi kt}{T}\right) \right.$$

$$\left. + \left[\frac{2}{T} \int_{t - T/2}^{t + T/2} g(x) \sin\left(\frac{2\pi kx}{T}\right) dx \right] \sin\left(\frac{2\pi kt}{T}\right) \right\} \qquad (5\text{-}7a)$$

The second and third terms of Equation (5-7a) can be combined using a trigonometric identity:

$$g(t) = \frac{1}{T} \int_{t - T/2}^{t + T/2} g(x)dx + \frac{2}{T} \sum_{k=1}^{\infty} \int_{t - T/2}^{t + T/2} g(x) \cos \left[\frac{2\pi k}{T} \quad (t - x) \right] dx. \qquad (5\text{-}7b)$$

This equation can then be rewritten in terms of convolution notation:

$$g(t) = g(t)* \left(\frac{1}{T} \; [U(t - T/2) \; - \; U\,(t + T/2)] \right)$$

$$+ \sum_{k=1}^{\infty} g(t)* \left(\frac{2}{T} \; [U(t - T/2) - U(t + T/2)] \cos\left(\frac{2\pi kt}{T}\right) \right), \qquad (5\text{-}7c)$$

where * indicates convolution.

The filter analysis system, which is equivalent to this time domain analysis-synthesis system, can be obtained by taking the Fourier transform of both sides of (5-7c) and making use of the convolution theorem:[34]

$$F\,[g(t)] = G(j\omega) = G(j\omega)\,\frac{\sin(\omega T/2)}{\omega T/2}$$

$$+ \sum_{k=1}^{\infty} G(j\omega)\left(\frac{\sin[(\omega + 2\pi k/T)T/2]}{(\omega + 2\pi k/T)T/2} + \frac{\sin[(\omega - 2\pi k/T)T/2]}{(\omega - 2\pi k/T)T/2}\right) \qquad (5\text{-}8)$$

Hence, the amplitude $c_k(t)$ of the kth harmonic appears at the output of the kth band pass filter (one of an infinite set of arithmetically spaced band pass filters) whose response is given by

$$H_k(j\omega) = \frac{\sin[(\omega + 2\pi k/T)T/2]}{(\omega + 2\pi k/T)T/2} + \frac{\sin[(\omega - 2\pi k/T)T/2]}{(\omega - 2\pi k/T)T/2} \qquad (5\text{-}9)$$

since by Equation (5-7a) the time domain output corresponding to that of the kth filter is of the form,

$$a_k(t)\cos\left(\frac{2\pi kt}{T}\right) + b_k(t)\sin\left(\frac{2\pi kt}{T}\right) = c_k(t)\cos\left[\frac{2\pi kt}{T} - \phi_k(t)\right].$$

The response of the kth filter is zero at all integral values other than k times the fundamental frequency, meaning that steady state harmonics other than the kth will not affect its output. However, if the kth harmonic and its neighboring harmonic frequencies are transient in nature, their Fourier transforms will not be concentrated strictly in lines and filter "cross-talk" will occur. With regard to the output of filter k, this interchange will result in the loss of some detail of the original $\bar{c}_k(t)$ plus the introduction of other frequencies which will appear as ripple on the graph of $c_k(t)$.

---

[34] Cheng, D.K.: Analysis of Linear Systems, Addison-Wesley, Reading, Mass., 1959, p. 227.

This equivalent filter system is useful as a theoretical tool for estimating the nature of the Fourier series representation of transient signals. In a real filter system of this type it would be impossible to extract the envelopes of the filter outputs without introducing another averaging process.

It is also useful to consider that the computation of the coefficients as defined by Equation (5-1) can be divided into two steps: (1) Hetrodyning to DC and (2) Averaging or low pass filtering. (Calculation of the amplitude and phase by Equation (5-2c) represents a third step in the computation which is not really necessary to define a harmonic tone.) The hetrodyning operation is perfectly straight forward: It simply involves the analog multiplication of the signal g(t) by quadrature pairs of harmonics tuned to each desired harmonic frequency of this signal. On the other hand, the integral averaging operation given by Equation (5-1) is not necessarily the best possible since the rejection of ripple due to transient envelopes is not optimum, and a precise adjustment of T is required to reject this ripple adequately even in the steady state; this becomes even more of a problem if the fundamental frequency of the signal is not constant.

5.4 Rectangular Filter Approach

Another analysis procedure which gives a "complete" representation of a quasi-periodic signal can be derived from the partitioning of the Fourier transform (frequency domain representation) of the signal into equal bands such that each harmonic frequency lies precisely in the middle of a corresponding band.

The operation of partitioning is equivalent to passing the signal through a set of band pass filters of the following form:

$$H_o(j\omega) = \begin{cases} 1, & |\omega| \leq \frac{\pi}{T} \\ \cdot & \\ 0, & \text{elsewhere} \end{cases}$$

$$H_k(j\omega) = \begin{cases} 1, & (k - \frac{1}{2})\frac{2\pi}{T} \leq |\omega| \leq (k + \frac{1}{2})\frac{2\pi}{T} \\ 0, & \text{elsewhere} \\ k = & 1, 2, 3, \ldots \end{cases} \tag{5-10}$$

Since this set of filters completely spans the frequency domain, the signal transform can be expressed by the following identity:

$$G(j\omega) = \sum_{k=0}^{\infty} H_k(j\omega)G(j\omega) \tag{5-11}$$

Then, taking the inverse Fourier transform of both sides we obtain:

$$g(t) = \sum_{k=0}^{\infty} F^{-1}[H_k(j\omega)]*g(t) \tag{5-12a}$$

$$= g(t)*\int_{-\pi/T}^{\pi/T} e^{j\omega t}\frac{d\omega}{2\pi} + \sum_{k=1}^{\infty} g(t)*\left\{\int_{(k-\frac{1}{2})\frac{2\pi}{T}}^{(k+\frac{1}{2})\frac{2\pi}{T}} e^{j\omega t}\frac{d\omega}{2\pi} + \int_{-(k+\frac{1}{2})\frac{2\pi}{T}}^{-(k-\frac{1}{2})\frac{2\pi}{T}} e^{j\omega t}\frac{d\omega}{2\pi}\right\} \tag{5-12b}$$

$$= g(t)*[\frac{1}{\pi t}\sin(\frac{\pi t}{T})] + \sum_{k=1}^{\infty} g(t)*[\frac{2}{\pi t}\sin(\frac{\pi t}{T})\cos(\frac{2\pi k t}{T})] \tag{5-12c}$$

$$= \frac{1}{T}\int_{-\infty}^{\infty} g(x)\,\frac{\sin[\frac{\pi(t-x)}{T}]}{\frac{\pi(t-x)}{T}}dx + \sum_{k=1}^{\infty}\frac{2}{T}\int_{-\infty}^{\infty} g(x)\cos[\frac{2\pi k(t-x)}{T}]\,\frac{\sin[\frac{\pi(t-x)}{T}]}{\frac{\pi(t-x)}{T}}dx \tag{5-12d}$$

$$= \frac{1}{T}\int_{-\infty}^{\infty} g(x)\,\frac{\sin[\frac{\pi(t-x)}{T}]}{\frac{\pi(t-x)}{T}}\,dx + \sum_{k=1}^{\infty}\left[\frac{2}{T}\int_{-\infty}^{\infty} g(x)\cos\frac{2\pi kx}{T}\,\frac{\sin[\frac{\pi(t-x)}{T}]}{\frac{\pi(t-x)}{T}}\,dx\right]\cos(\frac{2\pi kt}{T})$$

$$+ \sum_{k=1}^{\infty}\left[\frac{2}{T}\int_{-\infty}^{\infty} g(x)\sin(\frac{2\pi kx}{T})\,\frac{\sin[\frac{\pi(t-x)}{T}]}{\frac{\pi(t-x)}{T}}\,dx\right]\sin(\frac{2\pi kt}{T}) \qquad (5\text{-}12e)$$

Note that Equation (5-12e) is precisely the same form as (5-7a) except that the averaging operation $\int_{t-T/2}^{t+T/2}$ has been replaced by $\int_{-\infty}^{\infty} \frac{\sin[\frac{\pi(t-x)}{T}]}{\frac{\pi(t-x)}{T}}$, which corresponds to an ideal sharp-cutoff low pass filter. With this method of averaging, equal weighting is given to retaining the original envelope of the particular harmonic of interest and to excluding the ripple caused by other harmonics which may be highly transient. Moreover, in the steady-state portions of a signal which is to be harmonically represented, the ripple is completely rejected even if the fundamental period T is chosen incorrectly or is changing at a rate slow compared to 1/T and within a range small compared to 1/T.

Two examples of averaging operations which yield "complete" representations of transient signals have been presented. There may exist other such integral operations since it is only necessary that the integral weight function w(t-x) (considering that in all cases the integral limits are from $-\infty$ to ∞) must satisfy the identity

$$1 = \frac{1}{T}\,F[w(t)] + \frac{2}{T}\sum_{k=1}^{\infty} F[w(t)\cos(\frac{2\pi kt}{T})]. \qquad (5\text{-}13)$$

If a harmonic simulation of a quasi-periodic signal is desired, completeness is an important consideration since we cannot hope to resynthesize an original signal exactly by means of the $a_k(t)$ and $b_k(t)$ coefficients if some information about the signal is lost through their computation. However, it is also essential that the amplitudes and phases of the harmonics as calculated by Equation (5-2c) be interpretable. In general, this means that the spectral content of these derived "parameter-signals" should be substantially massed at frequencies much lower than 1/T.

5.5 Comparison of Three Fourier Series Representations of a Simple Test Signal

As mentioned in a previous section, the first step in computing the Fourier series coefficients involves hetrodyning the signal with quadrature pairs of the corresponding harmonic frequencies. If we suppose that the signal has the form given by Equation (5-6) (except that the phases are constant), the computation of the coefficient $a_k(t)$ first involves the analog multiplication of $g(t)$ with $2\cos(2\pi kt/T)$. This gives the result:

$$2g(t)\cos(2\pi kt/T) = \bar{c}_k(t)\cos(\phi_k) + [\bar{c}_{k-1}(t) + \bar{c}_{k+1}(t)]\cos(\phi_k)\cos(2\pi t/T)$$

$$+ [\bar{c}_{k-1}(t) - \bar{c}_{k+1}(t)]\cos(\phi_k)\sin(2\pi t/T) + \ldots \qquad (5\text{-}14)$$

The second step is to average, or equivalently, to low-pass-filter the hetrodyned signal. Ideally, this operation would extract the term $\bar{c}_k(t)\cos(\phi_k)$ and completely reject the other higher frequency terms. However, since the spectral characteristics of the harmonic envelopes are normally not band-limited, the Fourier transforms of the various terms of Equation (5-14) will overlap and no linear filter will be able to discriminate between them perfectly.

With regard to the choice of the optimum transfer characteristic for this low pass filter, we are presented with an infinite array of possibilities, and simultaneously we are confronted with the so-called acoustical "Uncertainty Principle"[35] which limits the accuracy with which amplitude and phase information pertaining to narrow band transient signals can be extracted. In order to provide a realistic illustration of this problem, the responses of three different averaging operators, each of which have quite simple time and frequency domain representations, to the test signal

$$s(t) = U(t)e^{-t/T}\cos(k2\pi t),\ k = 0, 1, 2, \ldots \tag{5-15}$$

will be compared. For simplicity we set T equal to unity. The response to sums of inputs of this type is obtained by superposition.

In the time domain, the three integral operators are defined by their responses to an arbitrary input g(t):

$$\text{(a)} \quad \int_{t-\frac{1}{2}}^{t+\frac{1}{2}} g(x)dx\ ,$$

$$\text{(b)} \quad \int_{t-1}^{t+1} g(x)\cos^2[\tfrac{1}{2}\pi(t-x)]dx\ ,$$

$$\text{(c)} \quad \int_{-\infty}^{\infty} g(x)\,\frac{\sin[\pi(t-x)]}{\pi(t-x)}\,dx\ . \tag{5-16}$$

These operators are equivalent to the following frequency domain filter response characteristics:

$$\text{(a)}\ \frac{\sin(\frac{1}{2}\omega)}{\frac{1}{2}\omega}\ , \qquad \text{(b)}\ \frac{\sin(\omega)}{\omega[1-(\omega/\pi)^2]}\ , \qquad \text{(c)}\ \begin{cases} 1, & |\omega| \leq \pi \\ 0, & \text{elsewhere.} \end{cases} \tag{5-17}$$

---

35. Kock, W.E. : "On the Principle of Uncertainty in Sound," J. Acoust. Soc. Am., 7(1), 56-58, 1935.

These filter response characteristics are shown graphically by Figure 5.1 (a), (b), (c), respectively. All three share the property that their response is zero for integral values of 2π. Type (a) corresponds to the most simple linear operation, integration over a single period. Type (b) corresponds to an integration over two periods with a $\cos^2$ weight function, featuring much reduced response to frequencies exceeding 2π radians per second. Type (c) requires a much more difficult time domain integration (the limits are infinite and the weight function is bounded only by C/x); however, the frequency response is perfectly flat below $\omega = \pi$ rps, whereas all frequencies above this value are rejected.

The ideal but absolutely unobtainable time domain response to s(t) (Equation (5-15)) is

$$r(t) = \begin{cases} U(t)e^{-t/\tau}, & k = 0 \\ 0, & k > 0. \end{cases} \tag{5-18}$$

By substituting s(x) for g(x) in Equation (5-16) it can be shown that the response of operators (a) and (b) (with $t > \frac{1}{2}$, $t > 1$, respectively) to s(t) both have the form

$$r(t) = [A(k,\tau)\cos(k2\pi t) + B(k,\tau)\sin(k2\pi t)]\, e^{-t/\tau}, \tag{5-19}$$

where the coefficients A and B for the (a) and (b) cases are given by

$$A_a(k,\tau) = \frac{2\tau\sinh(1/2\tau)}{1 + (2k\pi\tau)^2}, \qquad B_a(k,\tau) = \frac{-4k\pi\tau^2\sinh(1/2\tau)}{1 + (2k\pi\tau)^2}, \tag{5-20}$$

$$A_b(k,\tau) = \tau\sinh(1/\tau)\left[\frac{1}{1 + (2k\pi\tau)^2} - \frac{\frac{1}{2}}{1 + ((2k+1)\pi\tau)^2} - \frac{\frac{1}{2}}{1 + ((2k-1)\pi\tau)^2}\right]$$

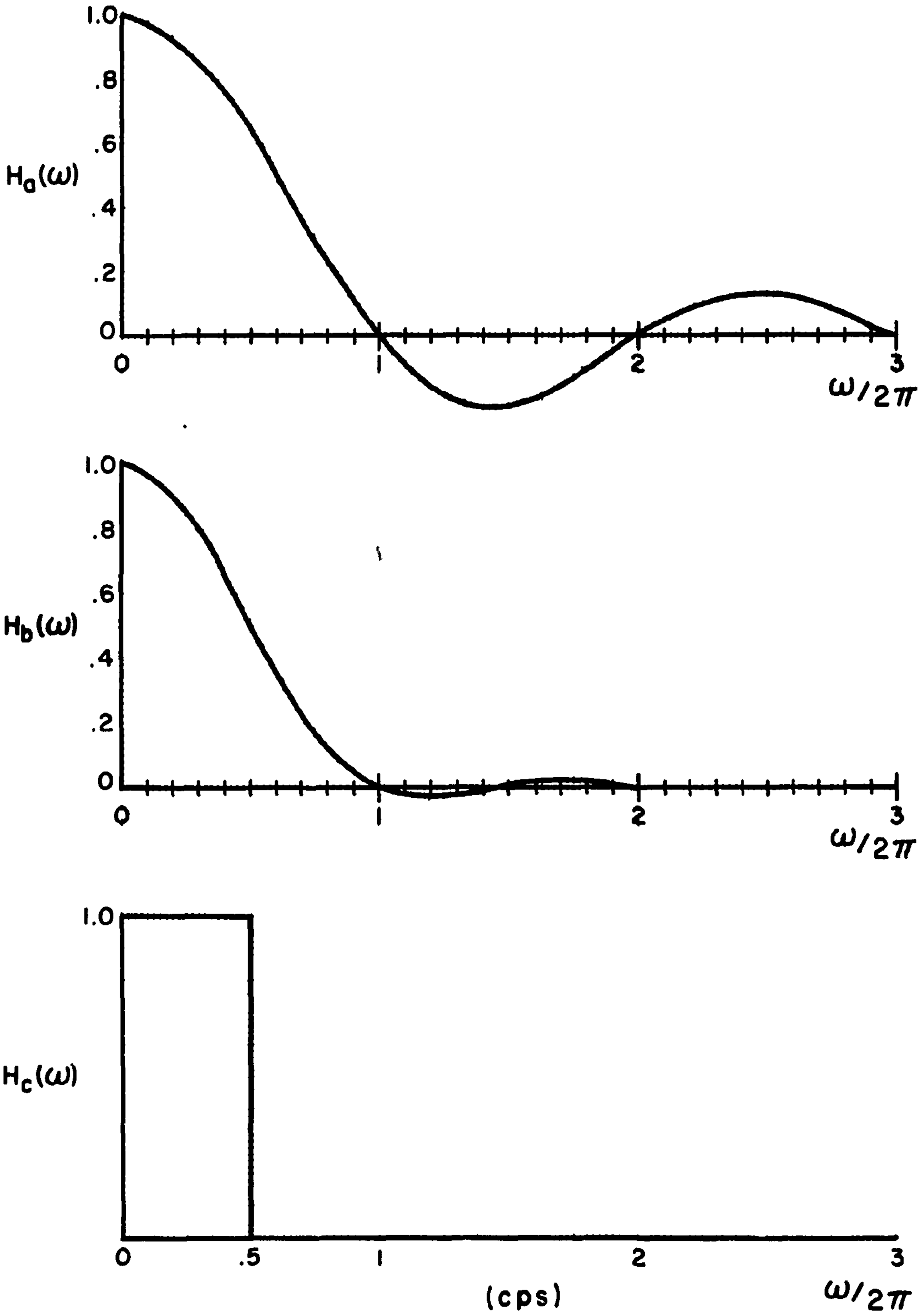


Figure 5.1 Low Pass Filter Characteristics Corresponding to Three Averaging Operators.

$$B_b(k,\tau) = -\tau\sinh(1/\tau)\left[\frac{2k\pi\tau}{1 + (2k\pi\tau)^2} - \frac{(k + \frac{1}{2})\pi\tau}{1 + ((2k+1)\pi\tau)^2} - \frac{(k - \frac{1}{2})\pi\tau}{1 + ((2k-1)\pi\tau)^2}\right]$$

Graphs of these coefficients for $k = 0$, 1, and 2 are given in Figure 5.2. These should be compared to the ideal case where $A(k,\tau)$ is equal to unity for $k = 0$ but is zero for nonzero values of k, and $B(k,\tau)$ is always zero. These curves together with Equation (5-14) allow an estimation of the "cross-talk ripple" which is superimposed on the graph of the derived Fourier coefficient, assuming that the coefficients of Equation (5-6) are behaving in an exponential manner. It should be remembered that the amplitude of a transient component to be rejected may be much greater than that of the DC component (after hetrodyning) which is to be extracted. Therefore, it is important that the rejection of non-DC transient components be as strong as possible. In both cases, with respect to the first harmonic ripple components, B is the least attenuated term. However, the type (b) response offers at least a 3 to 1 or 10 db improvement in attenuation of the term over the type (a) response for all values of τ. For the second harmonic ($k = 2$) and large values of τ the type (b) response offers a 4.8 to 1 improvement, i.e., an improvement of 13.6 db.

Operators (a) and (b) are both distinguished by finite integration limits. This is especially advantageous if the computation of the Fourier coefficients is to be done by digital computer. However, in the case of analog computation it becomes feasible to use filters with sharp-cutoff characteristics. This situation is epitomized by the frequency characteristic of operator (c). One might intuitively feel that this type of response would most effectively discriminate between low frequency and high frequency transient components. In actuality, there are limitations on the dynamic response of operator (c) to

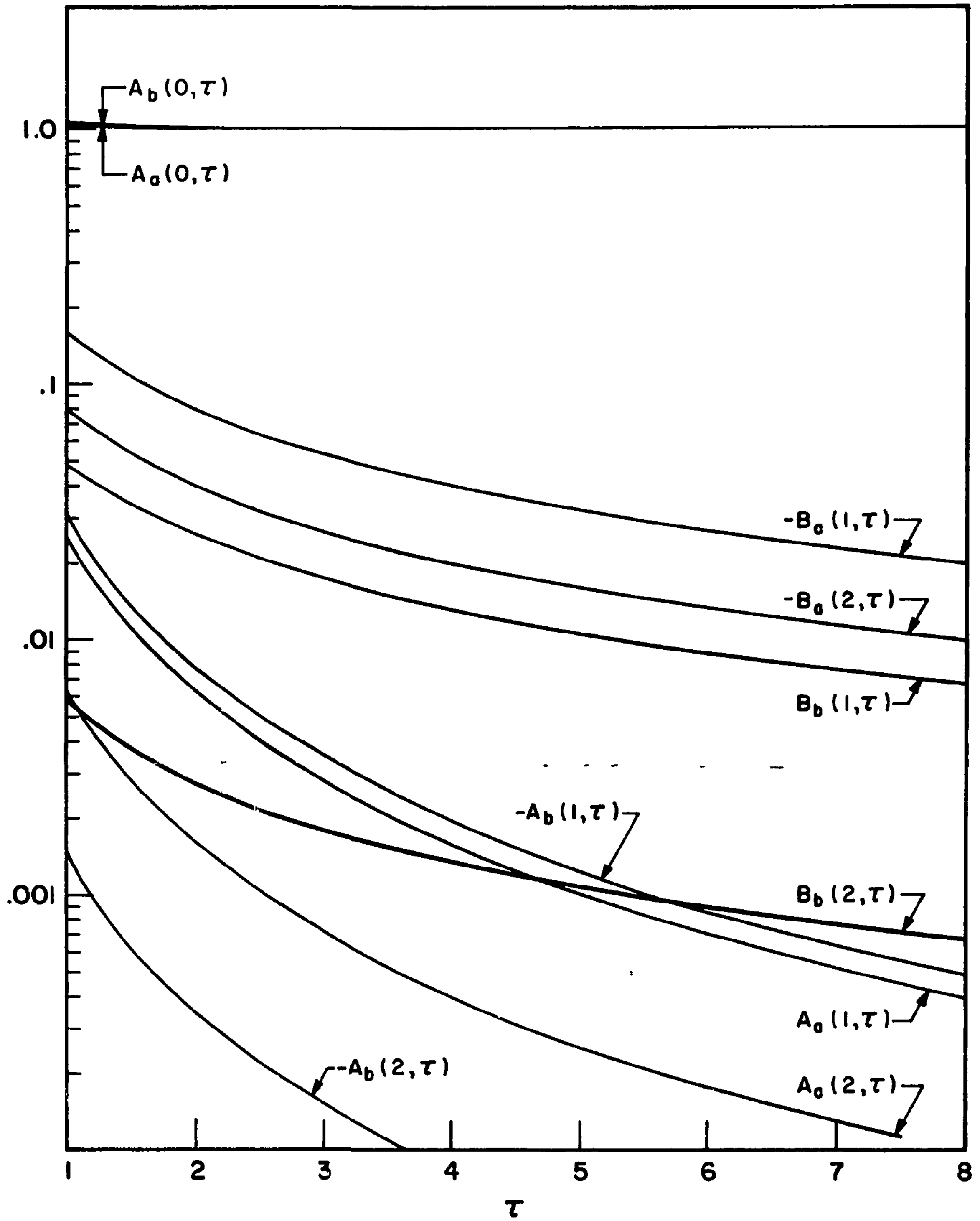


Figure 5.2 Coefficients of Equation (5-20) vs. Exponential Time Constant τ. (k is a Parameter)

rapid transients because of the slowly damped prediction and memory capabilities of the impulse response $\sin(\pi t)/\pi t$.

The sharp-cutoff filter response to the test signal s(t) can be calculated by substitution of s(x) given in Equation (5-15) for g(x) in Equation (5-16c):

$$\begin{aligned}
r(t) &= \int_{-\infty}^{\infty} s(x)\frac{\sin[\pi(t-x)]}{\pi(t-x)}\,dx = \int_{-\infty}^{\infty} s(t-x)\frac{\sin(\pi x)}{\pi x}\,dx \\
&= \int_{-\infty}^{t} e^{-(t-x)/\tau}\cos[k2\pi(t-x)]\frac{\sin(\pi x)}{\pi x}\,dx \\
&= e^{-t/\tau}\Bigg[\cos(k2\pi t)\bigg(\frac{1}{2\pi}\tan^{-1}[(2k+1)\pi\tau] - \frac{1}{2\pi}\tan^{-1}[(2k-1)\pi\tau] \\
&\qquad + \int_{0}^{t} e^{x/\tau}\cos(k2\pi x)\frac{\sin(\pi x)}{\pi x}\,dx\bigg) \\
&\qquad + \sin(k2\pi t)\bigg(\frac{1}{4\pi}\ln\Big[(1 + [(2k+1)\pi\tau]^2)/(1 + [(2k-1)\pi\tau]^2)\Big] \\
&\qquad + \int_{0}^{t} e^{x/\tau}\sin(k2\pi x)\,\frac{\sin(\pi x)}{\pi x}\,dx\bigg)\Bigg] \qquad (5\text{-}21)
\end{aligned}$$

For k = 0, this response reduces to:

$$r(t) = e^{-t/\tau}\left[\frac{1}{\pi}\tan^{-1}(\pi\tau) + \int_{0}^{t} e^{x/\tau}\,\frac{\sin(\pi x)}{\pi x}\,dx\right] \qquad (5\text{-}22)$$

Here again the form of Equation (5-21) is identical to that of Equation (5-19) except that in this case the A and B coefficients are dependent on time as well as the parameters k and τ. For example, when k is equal to zero, the integral term given in Equation (5-22) is not bounded with respect to increasing time, although r(t) is. The integral term introduces a rather disagreeable error which has a maximum at every odd integer value of t and a minimum at every even integer value. In other words, there is a relatively strong ripple component with frequency $\omega = \pi$ (the frequency of sharp cutoff) which does not decay as rapidly as the desired component $e^{-t/\tau}$.

For this more complex situation, compared to that of operators (a) and (b), it is necessary to use some sort of average criterion for comparing the errors of extraction or rejection of the envelopes for different values of k and τ. One convenient formulation is

$$\text{error} = \frac{\sqrt{\int_{.5}^{\infty} [s(t) - r(t)]^2 dt}}{\sqrt{\int_{.5}^{\infty} s^2(t) dt}}, \quad k = 0: \qquad (5\text{-}23a)$$

$$\text{error} = \sqrt{\int_{.5}^{\infty} r^2(t) dt}, \quad k \neq 0. \qquad (5\text{-}23b)$$

Values of error for operator (c) were calculated with the University of Illinois IBM 7094 computer as formulated by Equation (5-23) and are shown plotted as a function of τ for k = 0, 1, 2 in Figure 5.3. It is interesting that for k = 0, Equation (5-23a) as applied to the response of operators (a) and (b) to the test signal yields simply |A-1| as the error.

In summary, it appears that neither the rectangular time domain response operator (a) nor the rectangular frequency domain response operator (c) is optimal for evaluating the envelopes of frequency components which are rapidly changing over a large dynamic range, which is a reasonable description of parts of the envelopes of most musical tones. The operator (b) with its modified weight function represents a compromise between these two which is clearly better than either.

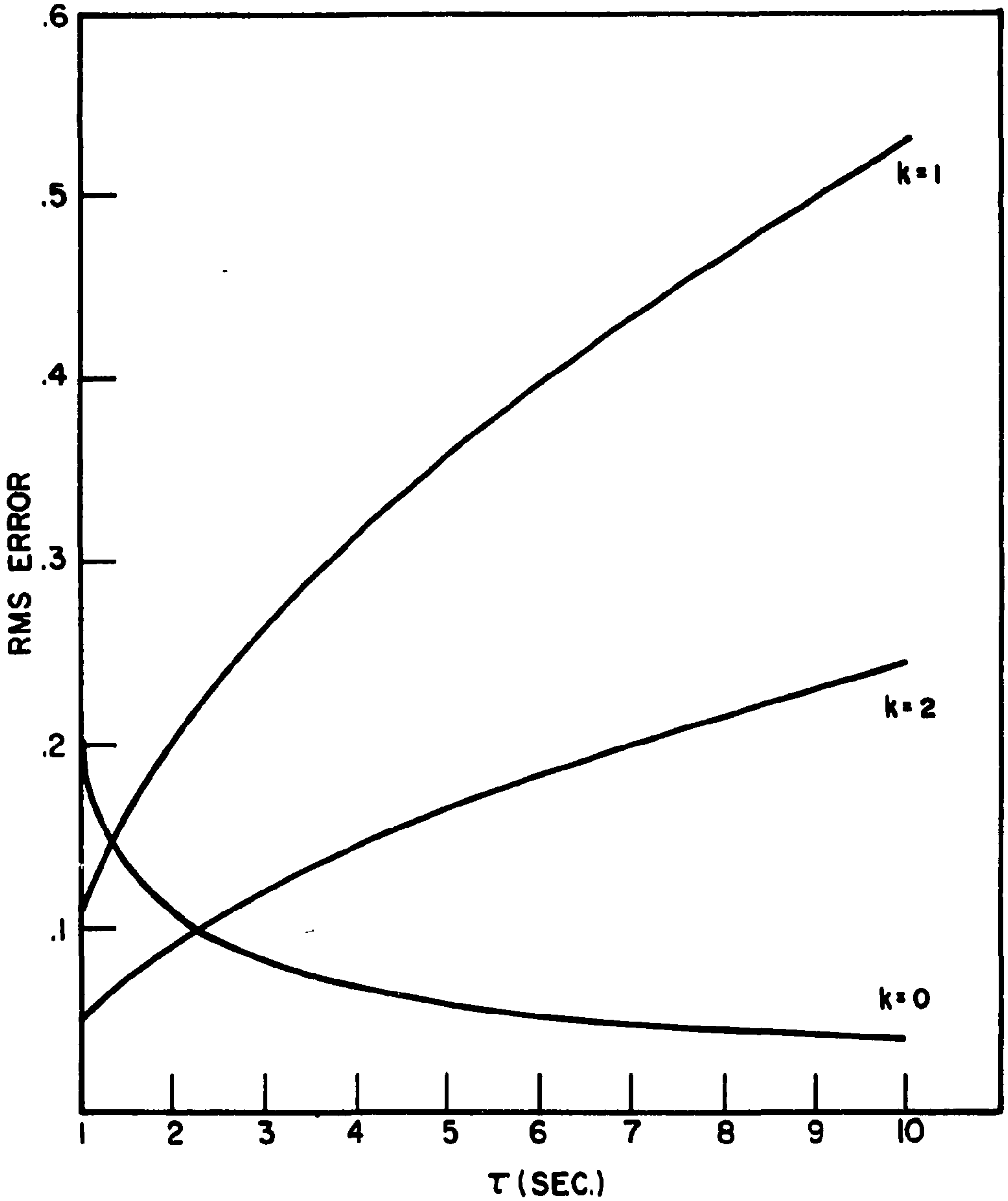


Figure 5.3 Rms. Error of Operation (c) Response to $e^{-t/\tau} \cos(k2\pi t)$ vs. τ. (k is a Parameter)

Of course, the "best operator response characteristic would depend on the particular test signal chosen. Equation (5-15) was selected because it is typical of the decay and attack curves found in music [s(t) may be thought of as a "decay", whereas s(-t) can be considered to be an "attack"], and it is relatively easy to handle on an analytical basis.

5.6 A Projected Analog System for Analysis and Resynthesis of Musical Tones

It is possible to construct by modification of the Harmonic Tone Generator design concept an analog analyzer for the calculation of instantaneous amplitudes and phases. The computational process consists of three steps: (1) Hetrodyning the signal with two quadrature components of each harmonic frequency, (2) separate low pass filtering of the hetrodyned components, and (3) computation of the amplitudes and phases (which amounts to the solution of right triangles). Figure 5.4 indicates how this might be done using a variable and fixed carrier frequency. Specifically, if the signal to be analyzed, s(t), has the form of Equation (5-6), the output of the electronic switch driven by the fixed frequency oscillator is

$$\sum_{k=1}^{n} \bar{c}_k(t)\cos\left[\frac{2\pi kt}{T} + \bar{\phi}_k(t)\right] \left[\frac{1}{2} + \sum_{j=1}^{\infty} \frac{\sin(j\pi a)}{j\pi}\cos(j2\pi f_c t)\right]$$

$$= \frac{1}{2} \sum_{k=1}^{n} \bar{c}_k(t)\cos\left[\frac{2\pi kt}{T} + \bar{\phi}_k(t)\right] + \sum_{j=1}^{\infty} \sum_{k=1}^{n} \bar{c}_k(t)\frac{\sin(j\pi a)}{j\pi}\cos\left[\frac{2\pi kt}{T} + \bar{\phi}_k(t)\right]\cos(j2\pi f_c t)$$

$$= \frac{1}{2} \sum_{k=1}^{n} \bar{c}_k(t)\cos\left[\frac{2\pi kt}{T} + \bar{\phi}_k(t)\right]$$

$$+\frac{1}{2} \sum_{j=1}^{\infty} \sum_{k=1}^{n} \bar{c}_k(t)\frac{\sin(j\pi a)}{j\pi}\left(\cos\left[2\pi\left(jf_c+\frac{k}{T}\right)t + \bar{\phi}_k(t)\right] + \cos\left[2\pi\left(jf_c-\frac{k}{T}\right)t - \bar{\phi}_k(t)\right]\right). \qquad (5\text{-}24)$$

where a/f_c is the pulse duration.

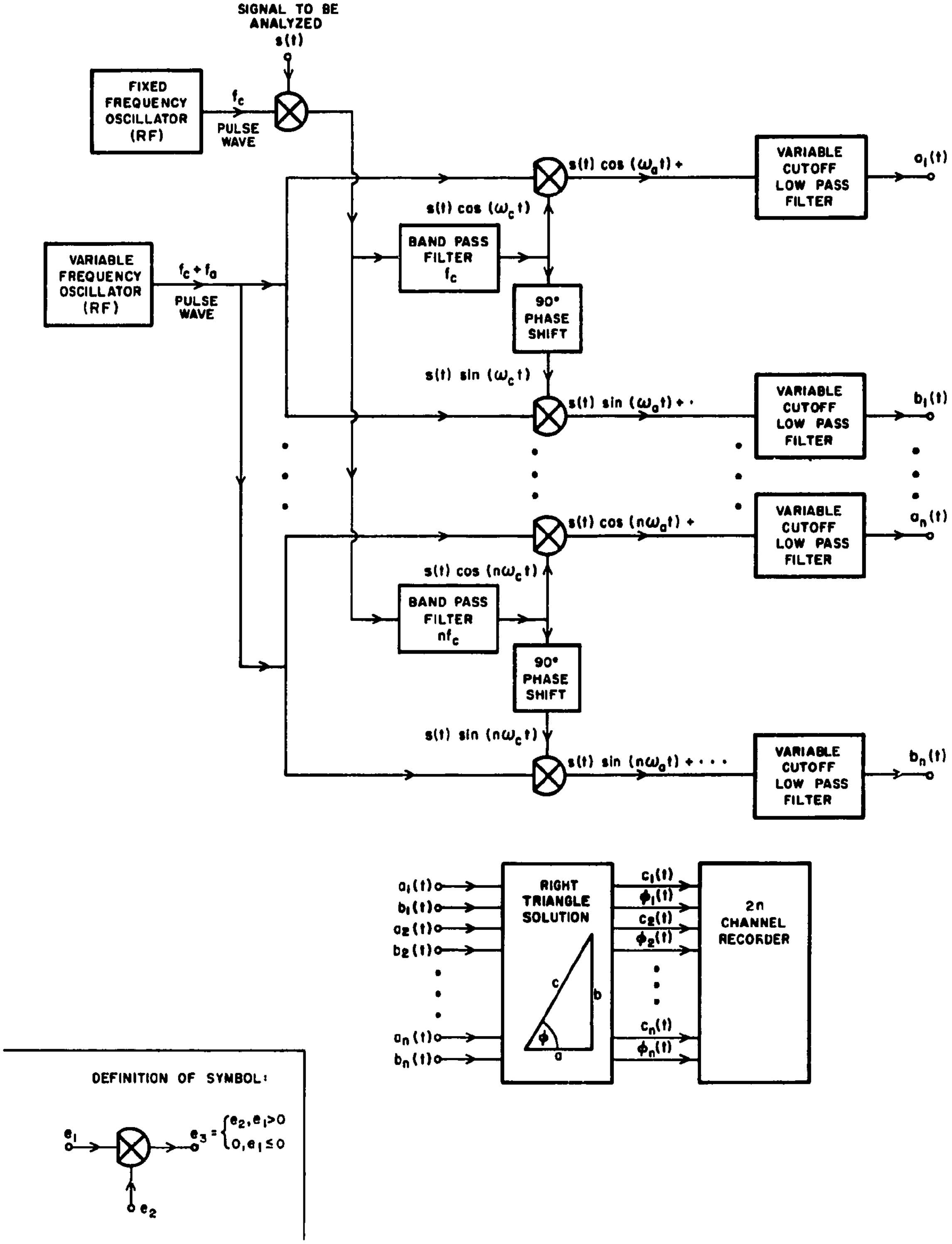


Figure 5.4 Analog System for Analysis of a Transient Harmonic Tone.

The band pass filter for extracting the jth harmonic should be designed to pass the frequencies $jf_c - (j+1)/T$ to $jf_c + (j+1)/T$ without distortion, i.e., the amplitude response should be flat and the phase response linear over this frequency range. However, frequencies below $(n+1)/T$ and above $(j+1)f_c - (n+1)/T$ should be rejected. Except for some distortion of the harmonics exceeding the jth which are not of interest, the output of the band pass filter within a constant factor is $s(t)\cos(2j\pi f_c t)$. The output of the 90° phase shift circuit is $s(t)\sin(j2\pi f_c t)$. The output of one of the electronic switches driven by the variable frequency oscillator, which is set to $f_c + 1/T$, is

$$s(t)\cos(j2\pi f_c t)\left[\frac{1}{2} + \sum_{i=1}^{\infty} \frac{\sin(i\pi a)}{i\pi} \cos[i2\pi(f_c + 1/T)]\right]$$

$$= \frac{\sin(j\pi a)}{j\pi}\, s(t)\cos(j2\pi t/T) + \text{high frequency terms.} \tag{5-25}$$

The variable cutoff low pass filter removes the high frequency terms and averages the hetrodyned low frequency component.

The advantages of this method are that any fundamental frequency can be handled and the tone may be played directly into the analyzer without an intermediate tape recorder. Also, variable cutoff low pass filters are commercially available which would allow adjustment of the bandwidth for the best compromise between retaining the transient characteristics of the Fourier coefficients and rejecting ripple. The amplitude outputs could be used as inputs to the Harmonic Tone Generator, which would in turn resynthesize the original tone; the two units connected together would function in effect as a musical tone vocoder. Because all harmonics are analyzed simultaneously, the relative phases of the harmonics are correct with respect to one another, although the stability of the absolute phase depends on the stability of the variable and fixed carrier frequencies.

The obvious drawbacks of this system, of course, are its complexity and its cost. It requires separate fixed band pass filters, low pass filters, complex nonlinear networks, and a multi-channel chart recorder. The system is tremendously simplified if only one frequency at a time is analyzed or if the fundamental frequency is fixed. For example, Figure 5.5 shows a system for measuring the amplitude and phase derivative of a particular frequency f_a (the absolute or relative phases are of no consequence since the analyses of different frequencies are not performed simultaneously). Note that some simplification was possible as compared to a single harmonic channel of the system in Figure 5.4. A time mark must be recorded on the tape immediately preceding the start of the tone to be analyzed in order to identify the time reference of the analysis.

Figure 5.6 illustrates a system which analyzes at a standard fundamental frequency f_a. The band pass filter is fixed and its pass and rejection bands are optimized to pass one harmonic and reject the others. The filter bandwidth is very narrow compared to its center frequency; hence, the symmetry of the filter transfer response is arithmetic to a close approximation. The harmonic to be analyzed is selected by choosing the proper hetrodyne frequency. The amplitude of the output is the desired amplitude, and the phase derivative can be computed by means of a phase-detector and differentiator. Other fundamental frequencies can be accommodated by means of a variable speed tape recorder, although care must be taken to insure that proper equalization of the original

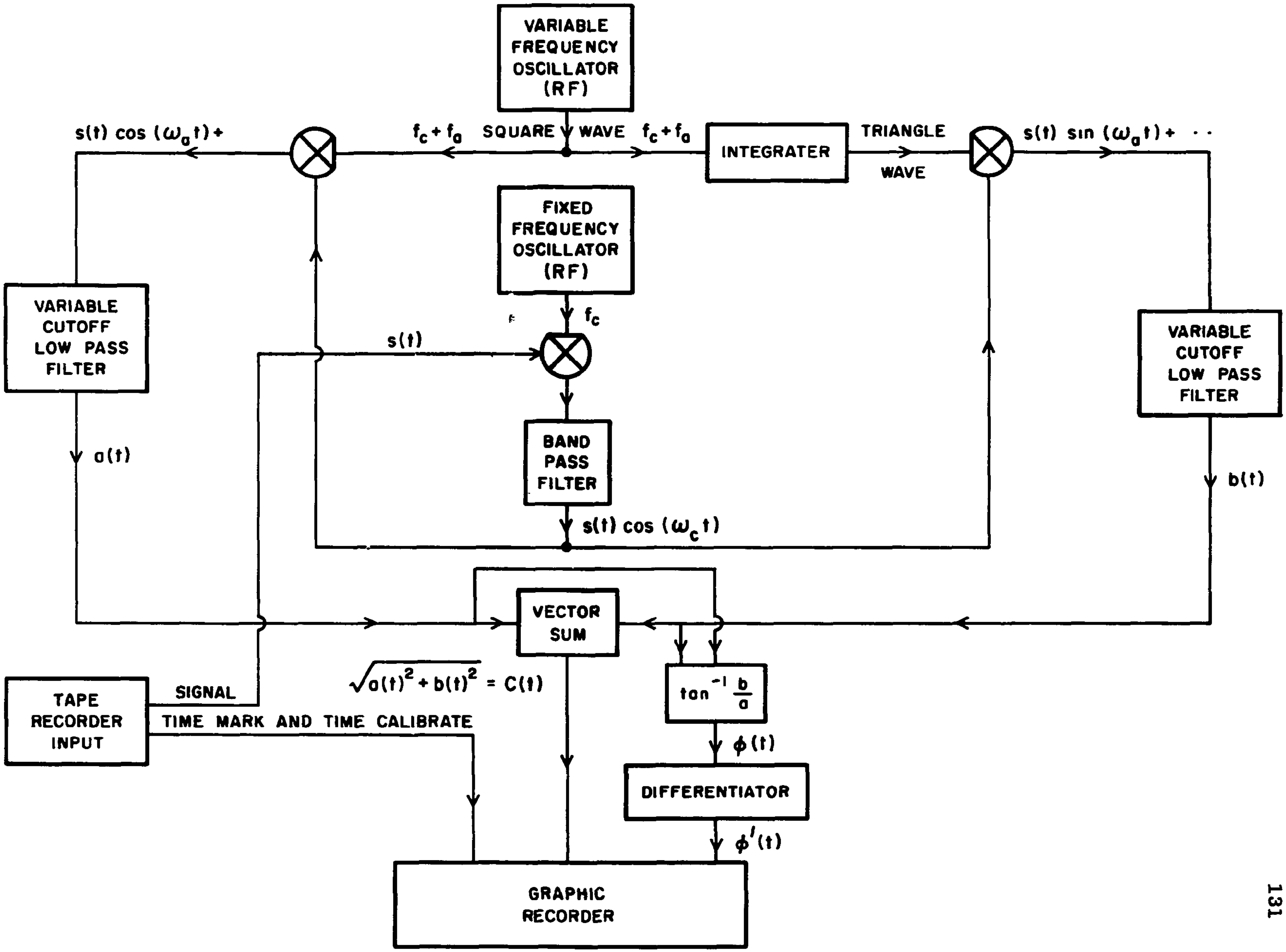


Figure 5.5 Single Variable Frequency Transient Tone Analyzer.

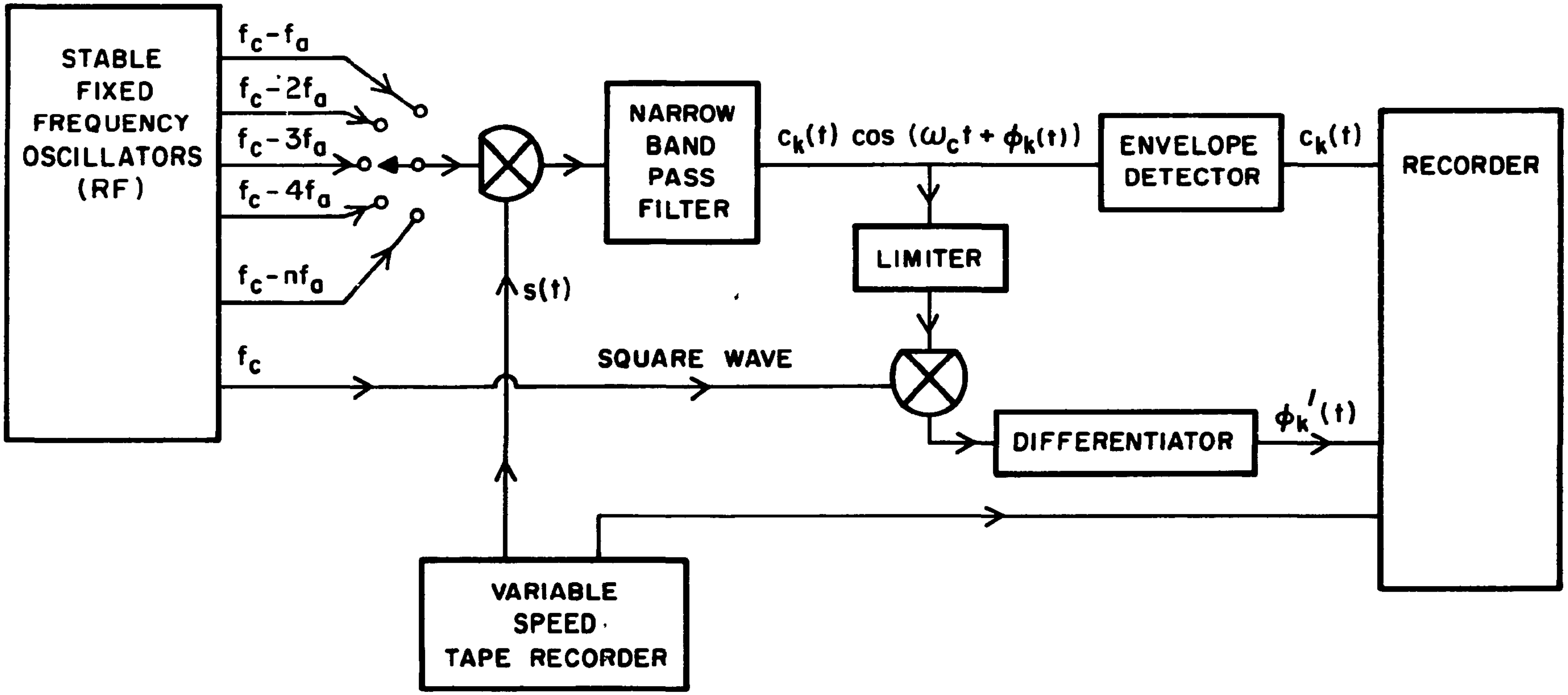


Figure 5.6 Single Fixed Frequency Transient Tone Analyzer.

frequencies is maintained, and that the speed is constant.

The chart speeds for plotting of musical instrument transients must be quite fast--perhaps too fast for ink types. More accurate ink curves could be obtained by recording the data at high speed on an FM tape recorder and playing back at reduced speed. Light-beam oscillographs are commercially available, but the graphs resulting from these units are generally inferior in quality to ink types, must be fixed by photographic processes, and are difficult to reproduce. Another possible solution is continuous photography of oscilloscope screen representations of the transients.

Chapter 6

A PERCEPTUAL CHARACTERISTIC OF HARMONIC TONES: THRESHOLD FOR "HEARING OUT" THE UPPERMOST PARTIAL

6.1 Introduction

When a person hears a harmonic tone in musical context he does not normally separate out the individual frequencies which the tone comprises. Rather he hears the tone as a single entity or "Gestalt". Nevertheless, it is a well-known fact that the listener can sometimes perceive the pitch of a harmonic as distinct from that of the tone itself if one or more of the following conditions are satisfied: (a) The harmonic tone has a relatively fixed pitch and amplitude structure; (b) the listener can form a proper mental image of the frequency he is to hear out in advance; (c) the harmonic in question is sufficiently intense relative to the rest of the tone (and especially relative to the partials adjacent to this harmonic). Limitations on this type of perception were recently investigated by Plomp.[36] Two of his conclusions are: (1) ". . . partials of a complex sound can be 'heard out' only if their frequency separation exceed the [so-called] critical bandwidth." Consequently, it is difficult to hear out harmonics above the ninth for tones with flat spectra. (2) ". . . the old conviction, that the third and fifth partials -- in general the uneven ones -- are easier distinguished than the even partials, is not confirmed [by Plomp's data]." (Several references to observers who do share this conviction are given in his article.)

---

[36] Plomp, R. . "The Ear as a Frequency Analyzer," J. Acoust. Soc. Am., 36(9), 1623-1636, 1964.

Plomp's results are based on a test which uses forced decision comparisons of pure sine tones to harmonic tones consisting of the first 12 partials at equal intensity. The experiment described here uses an entirely different method to determine the relative amplitude of a partial necessary for it to be heard out, and, incidentally, whether or not the odd partials are easier to hear separately than the even ones. In the latter case, the data derived from this experiment shows agreement with Plomp's conclusion.

6.2 Experimental Method

Fourteen listeners, who are highly trained musicians or have had considerable experience in sound perception, were tested separately to determine their individual limitations with respect to hearing out the individual partials of harmonic tones. At each of five fundamental frequencies 65, 130, 260, 520, and 1040 cps., a group of five tones was presented to the listeners via a loudspeaker (Acoustics Research AR-3). The five tones consisted successively of the first two partials, the first three partials, and so on. Each partial of a tone except the highest was set at the same subjective intensity level (in phons) as measured by a sound level meter (B&K 2203). The levels were set by the appropriate attenuator controls on the Harmonic Tone Generator, which was used as the sound source in this experiment. Since the experimental laboratory was not soundproof, it was necessary to use playback levels which were both "comfortable" and at the same time sufficiently intense to allow the listener to ignore any background noise present. These requirements resulted in the choices 70 phons ar 65, 260, and 520 cps., 80 phons at 130 cps., and 60 phons at 1040 cps.

Each listener was instructed to adjust the intensity of the highest harmonic of a given tone to the lowest level where without question he (or she) could distinctly perceive the pitch of this harmonic as separate from that of the tone itself. This level was measured in acoustical decibels with the sound level meter and a wave analyzer (Hewlett-Packard 302A) tuned to the perceived harmonic frequency. Later the value was converted to subjective intensity units. The wave analyzer was used to insure that the actual acoustical intensity of each harmonic was being measured, thus eliminating errors which could be introduced by the effects of system intermodulation distortion and by distortion of the original harmonic components. The experimental set-up is given by Figure 6.1. The sound level meter microphone was placed near the listener's left ear (each listener was instructed to stop his right ear) so that the reading would correspond as close as possible to the sound level at this ear.

Errors in the measurements were bound to occur because of a number of factors such as the variation in the room background noise, the variation of the room standing-wave pattern with different physical positions of the listener and the experimenter (an attempt was made to keep this constant), the variation of the listener's interpretation of the experimenter's definition of the level at which a harmonic is perceived separately, subjective harmonics, and the possibility of confusing imagination with perception, as well as ordinary measurement errors. Standing waves would not have been a problem had earphones of sufficient quality been available. However, all of the earphones tested had a tendency toward introducing excessive distortion and sympathetic vibration sounds, and provided an auditory "closeness" which greatly increased the level of auditory fatigue. In an attempt to average out the effects of fatigue from the beginning to the end of the experiment, the order of the fundamental frequencies of the tones presented was rotated.

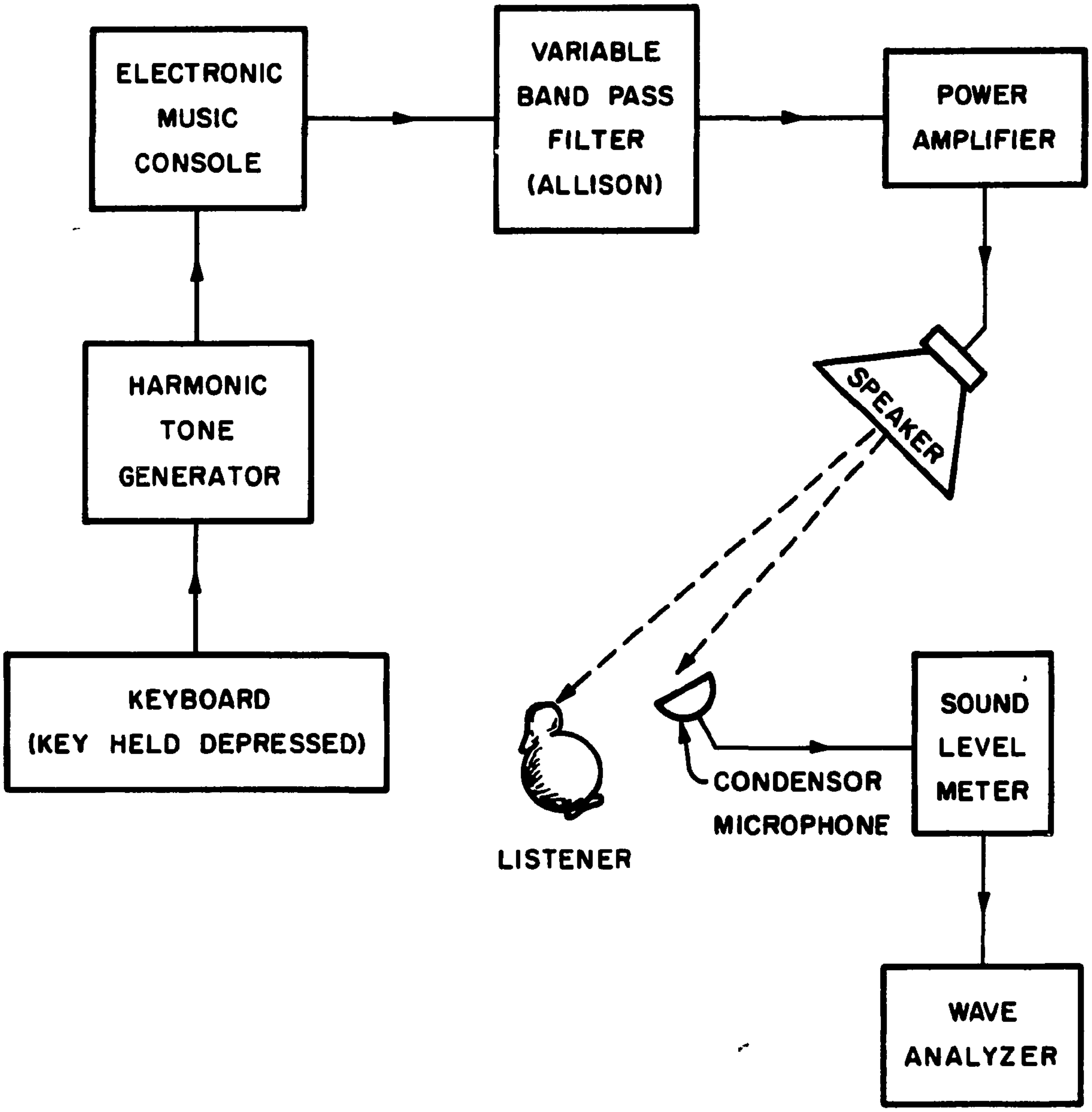


Figure 6.1 Test Configuration for the Psychoacoustic Experiment: "Hearing Out" the Uppermost Partial of a Harmonic Tone.

6.3 Calculations

For each of the harmonic tones presented the levels of the individual constituent harmonics are set to equal subjective intensity levels in phons. Since sound level meter measurements are made in acoustical decibels (db) it is necessary to translate from the specified level in phons to the level in dbs by means of the Fletcher-Munson loudness curve.[21] The loudness of a harmonic tone excepting the nth partial is calculated by the conventional rule of energy combination. That is, the total loudness L_T is given in terms of the individual loudnesses which are all equal to L_I by

$$L_T = 10\log_{10} \sum_{k=1}^{n-1} [10^{L_I/20}]^2 = 10\log_{10}[(n-1)10^{L_I/10}]$$

$$= L_I + 10\log_{10}(n-1) \qquad \text{(phons)} \qquad (6\text{-}1)$$

The threshold criterion of most interest to me was the difference between the loudness of the nth partial required for the listener to hear it out. This gives the relative subjective intensity threshold of the uppermost harmonic when this harmonic is between the second and the sixth for five fundamental frequencies. The data values so derived are plotted in Figure 6.2 for each of the five frequencies together with the mean threshold values and the average absolute deviations averaged over the 14 listeners. Each average absolute deviation was computed as the average of the absolute values of the differences between the corresponding data values and the mean value.

6.4 Conclusions

Because of the large spread of the data it is difficult to infer any strong general conclusion about the results other than that the average threshold value for hearing out the highest partial is about -20 phons,

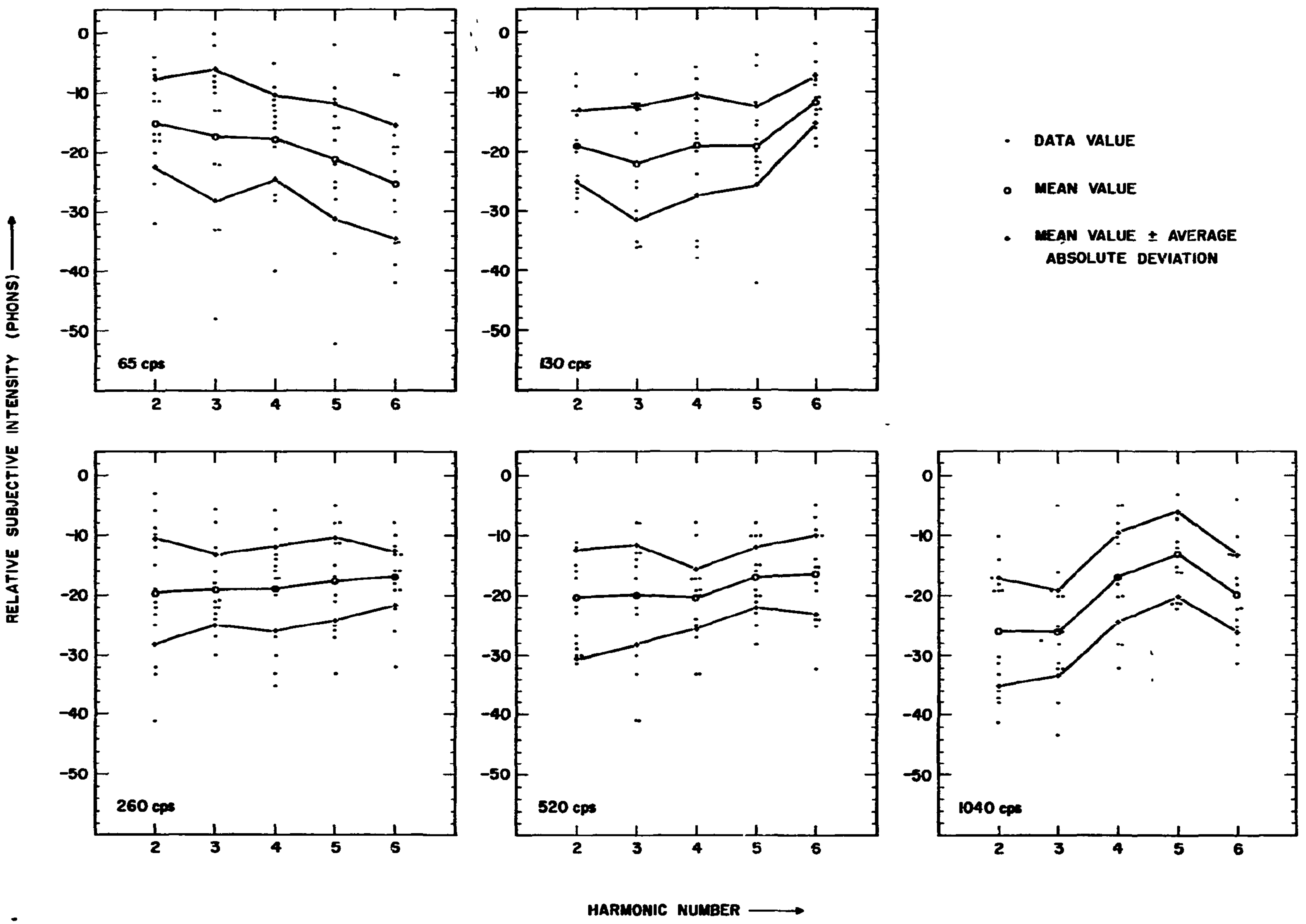


Figure 6.2 Threshold Level of Uppermost Partial: Test Results.

independent of frequency and harmonic number. The mean values show a drop of 10 phons from the second to the sixth harmonic for the 65 cps. tones and show a non-monotonic increase of 7 phons for the 130 cps tones. The 260 cps. and 520 cps. tones show only slight increases of 2 5 and 4 phons For the 1040 cps. tone the average threshold is a constant -26 phons for harmonics 2 and 3 but shows a sharp increase for higher harmonics. Three listeners could not hear out one or more of the fourth, fifth, and sixth partials of the 1040 cps. tones and the sixth partial of a 520 cps. tone under any available means for increasing the strength of these partials. This phenomenon of relative uncertainty in discriminating partials above the fourth for frequencies exceeding 500 cps. is in agreement with the results of Plomp's investigation. Only in the case of the 130 cps fundamental are the mean threshold values of the odd harmonics lower than they would be if they followed the general trend of change with respect to harmonic number. With this exception none of the data indicates that the odd harmonics are easier to hear out than the even ones under steady state conditions.

6.5 The Listeners

Participants in this listening experiment were:

1. Theodore Murname
2. Royal MacDonald
3. Bruno Nettl
4. Barbara Shapiro
5. David Schroer
6. Gary Grossman
7. Rhoda Grossman

8. Donald Andrus
9. Neely Bruce
10. James Beauchamp
11. Amy Larson
12. Thomas Howell
13. Julia Sentman
14. Burt Levy

None of the listeners experienced any difficulty in hearing out the partials, the intensity of which were under their own control. Some persons took much more time deliberating over the correct threshold values than others. The time for the complete testing of one listener ranged from 40 minutes to 70 minutes. Most people experienced some auditory fatigue because of the long tone durations used in the test. Some listeners were extremely acute at certain frequencies, but radically less so at other frequencies. The average thresholds for individuals ranged from 9 phons to 26 phons.

BIBLIOGRAPHY

1. Bachem, A. : "Chroma Fixation at the Ends of the Musical Scale," J. Acoust. Soc. Am., 20(5), 704-705, 1948.

2. Berger, F.B., MacRae, D.: "Mathematical Operations on Waveforms-II", Waveforms, Chap. 19, 679-686, M.I.T. Radiation Laboratory Series, Vol. 19.

3. Cheng, D.K.: Analysis of Linear Systems, Addison-Wesley, Reading, Mass., 1959, p. 227.

4. Churchhill, R V.: Fourier Series and Boundary Value Problems, McGraw-Hill, New York, 1941, pp. 65, 70.

5. Culver, C.A.: Musical Acoustics, McGraw-Hill, New York, 1956, p. 95.

6. Dorf, R.H.: Electronic Musical Instruments, Radio Magazines, Inc., Mineola, N Y., 1958, pp. 25-45.

7. Fletcher, H. and Munson, W.A.: "Loudness, Its Definition, Measurement, and Calculation," J. Acoust. Soc. Am., 5(2), 82-108, 1933.

8. Fletcher, H.: "The Pitch, Loudness and Quality of Musical Tones," Bell Telephone System Monograph, No. 1397, 1946.

9. Haynes, N.M.: Elements of Magnetic Tape Recording, Prentice-Hall, Englewood Cliffs, N.J., 1957, pp. 183-193.

10. Hiller, L.A., Jr.: "An Integrated Electronic Music Console," J. Audio Eng. Soc., 13(2), 142-150, 1965.

11. Hilsum, C.: "Multiplication by Semiconductors," Electronic Engineering, 30(369), 664-666, 1958.

12 Hunter, J.L.: Acoustics, Prentice-Hall, Englewood Cliffs, N J., 1957, pp. 253-256.

13. Knudsen, V.O.: "Modern Acoustics and Culture," University of California Press, Berkeley, California, 1937, pp. 22-31.

14. Kock, W.E.: "On the Principle of Uncertainty in Sound," J. Acoust. Soc. Am., 7(1), 56-58, 1935

15. Kraicer, A.: "Two Four-Quadrant Electronic Multipliers," Radio and Electronic Engineer, 25(1), 5-11, 1963

16. LeCaine, H.: "Electronic Music," Proc. Inst. Rad Eng., 44(4), 457-478, 1956

17. Lewis, D.: University of Iowa Studies in the Psychology of Music, Vol. IV, University of Iowa Publication, Iowa City, 1937, p. 346.

18. Licklider, J.C.: "Effects of Changes in the Phase Pattern Upon the Sound of a 16-Harmonic Tone," J. Acoust. Soc. Am., 29(6), 780 (abstract), 1957.

19. Luce, D A.: "Physical Correlates of Nonpercussive Musical Instrument Tones," unpublished Doctoral Dissertation, M.I.T., 1963.

20. Marshall, B.O.: "An Electronic Multiplier," Nature, 167, 29, 1951.

21. Mathes, R.C. and Miller, R.L.: "Phase Effects in Monaural Perception," J. Acoust. Soc. Am., 19(5), 780-797, 1947.

22. Mathews, M.V.: "The Digital Computer as a Musical Instrument," Science, 142 (3592), 553-557, 1963.

23. Miessner, B.F.: "Electronic Music and Instruments," Proc. Inst. Rad. Eng., 24(11), 1427-1463, 1936.

24. Moog, R.A.: "Voltage-Controlled Electronic Music Modules," J. Audio Eng. Soc., 13(3), 200-206, 1965.

25. Olmstead, J.M.: Real Variables, Appleton-Century-Crafts, New York, 1959 p. 283.

26. Olson, H.F.: Acoustical Engineering, D. Van Nostrand Co., New York, 1957, p. 610.

27. Pierce, J.R. and David, E.D.: Man's World of Sound, Doubleday & Co., Garden City, N.Y., 1958, p. 98.

28. Plomp, R: "The Ear as a Frequency Analyzer," J. Acoust. Soc. Am., 36(9), 1628-1636, 1964.

29. Schoeder, M.R.: "New Results Concerning Monaural Phase Sensitivity," J Acoust. Soc. Am., 31(11), 1579 (abstract), 1959.

30. Shower, E.G. and Biddulph, R.: "Differential Pitch Sensitivity of the Ear," J. Acoust. Soc. Am., 3(2), 275-287, 1931.

31. Ward, W.D.: "Subjective Musical Pitch," J. Acoust. Soc. Am., 26(3), 370-380, 1954.

32. Weldon, L A.: "Boost for Electronic Tuning: Part I," Electronics, Vol. 37, April 6, 1964, pp. 49-53.

33. Wilcox, R.H.: "Use of a Diode Ring as a Four Quadrant Multiplier," Review Scientific Instruments, 30(11), 1009-1011, 1959.

VITA

James Beauchamp was born on October 17, 1937 in Wayne County, Michigan. While attending Redford High School in Detroit, he studied music theory and composition and played trumpet in the school band and orchestra. After completing two years at Albion College, he began studies in electrical engineering at the University of Michigan in 1957. He received a B.S. in E.E. in 1960 at U. of M. and an M.S. in E.E. in 1961 from the same institution. From August, 1961 to August, 1962 he worked for Space Technology Laboratories, Inc. in Los Angeles. Since September, 1962, he has been the recipient of the Magnavox Fellowship in Electrical Engineering and Music at the University of Illinois. Along with his regular course work in electrical engineering he has engaged in research in electronic music at the University of Illinois Experimental Music Studio and completed several courses in the School of Music, including music composition.